University of Bath

PHD

An improved approach to data analysis and interpretation in transformer condition assessment based on unsupervised neutral network

Thang, Ka Fei

Award date:
2002

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 05. May. 2023

UNIVERSITY OF
BATH

AN IMPROVED APPROACH TO DATA ANALYSIS & INTERPRETATION IN TRANSFORMER CONDITION ASSESSMENT BASED ON UNSUPERVISED NEURAL NETWORK

Submitted by Ka Fei Thang
for the degree of PhD
of the University of Bath

2002

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

UMI Number: U601725

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U601725
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIV... BATH
... 20 NOV 2002
Ph.D.

ACKNOWLEDGEMENTS

I would like to thank Prof. R. K. Aggarwal for his invaluable support, guidance and advices during the interval of this research project and the preparation of this thesis. Working with him has been a pleasure due to his patience, kindness and reassurance throughout the challenging period of this research project.

I am also very grateful to Dr. Tony McGrail, Mr. David Esp, Dr. Dan Faircloth and Mr. Francis Waite of the National Grid Company, UK for their invaluable and unconditional support in financial resources, availability of data and expert advices throughout this research project. Particularly, I would like to acknowledge Dr. Tony McGrail and Mr. David Esp for their initial work in this project and their inspiration and dedication for the continuation and extension of their initial work in this research project.

I would also like to express my deepest appreciation to my family for their unreserved support and trust in me for continuing this research project. Without their support, I would not have endured the hardships and difficulties throughout the interval of my postgraduate study.

Finally, I gratefully acknowledge the University of Bath and the Committee of Vice-Chancellors and Principals (CVCP), UK for their financial support both in terms of tuition fees and living expenses throughout the interval of this research project.

ABSTRACT

This thesis describes the investigation and development of an improved and novel approach for the analysis and interpretation of the condition assessment data of power transformers with the ultimate aim of enhancing the condition assessment process; this is achieved by placing the main emphasis on the dissolved gas analysis (DGA) data of power transformers and their bushings. In addition, the proposed approach is also applied on a limited scale for the analysis of a set of on-line monitoring data comprising of various sensor measurements collected from a power transformer.

Through application of the proposed approach on the DGA data of power transformers and their bushings, it is clearly demonstrated that the self-organising map (SOM) is able to unearth the inherent characteristics of the DGA data and these revealed features can be presented in a discernible and comprehensible format. Importantly, it has also been verified that these revealed features can be associated with some conditions of power transformers and their bushings. The foregoing findings are advantageous due to the fact that the interpretation of the DGA data can be produced in a more accurate and convincing manner because it is based on the actual characteristics of the DGA data. More importantly, the approach developed could provide an improved means of monitoring the condition of power transformers and their bushings based on the "visualisation" of the movement of DGA trajectory.

Finally, it is shown that the application of SOM for analysis of the multiple sensor data collected from a power transformer also produces interesting results, in which the revealed features can be utilised for hourly on-line monitoring of the operating condition of power transformer. The foregoing means of on-line monitoring could potentially benefit the operator since more information can be obtained in this manner when compared to the conventional means of displaying the variation of sensor variables in time-sequence plots.

CONTENTS

INDEX TO FIGURES

INDEX TO TABLES

LIST OF ABBREVIATIONS

C_2H_2	Acetylene
ABM	Aged-Based Maintenance
AI	Artificial Intelligence
AQE	Average Quantisation Error
BMU	Best Matching Unit
CO_2	Carbon Dioxide
CO	Carbon Monoxide
CEGB	Central Electricity Generation Board
CIGRÉ	Conseil International des Grands Réseaux Électriques
CBM	Condition-Based Maintenance
FAN	Cooling Circuit 1 – Fans
PUMP	Cooling Circuit 2 – Pump
CT	Current Transformer
DM	Data Mining
DC	Direct Current
DGA	Dissolved Gas Analysis
EF	Electrical Fault
EOL	End-Of-Life
C_2H_6	Ethane
C_2H_4	Ethylene
EP	Evolutionary Programming
ES	Expert System
FRA	Frequency Response Analysis
FEP	Fuzzy Evolutionary Programming
FES	Fuzzy Expert System
FNN	Fuzzy Neural Network
GC	Gas Chromatograph
HPLC	High Performance Liquid Chromatography
HV	High Voltage

H_2	Hydrogen
IEEE	Institute of Electrical and Electronics Engineers
IEC	International Electrotechnical Commission
KDD	Knowledge Discovery in Databases
LTC	Load Tap Changer
LV	Low Voltage
CH_4	Methane
NGC	National Grid Company
NN	Neural Network
N_2	Nitrogen
NaN	Not-a-Number
ORS	Oil Relative Saturation
OTMS	Oil Temp. at Moisture Sensor
OLTC	On-Load Tap Changer
O_2	Oxygen
PPM	Part-Per-Million
PD	Partial Discharges
PC	Phase Current
RF	Radio Frequency
RFI	Radio Frequency Interference
RHA	Relative Humidity – Ambient
SOM	Self-Organising Map
SOPN	Self-Organising Polynomial Network
TA	Temperature – Ambient
TBO	Temperature – Bottom Oil
TLTC	Temperature – Load Tap Changer
TTO	Temperature – Top Oil
TF	Thermal Fault
TBM	Time-Based Maintenance
TE	Topographic Error
TCG	Total Combustible Gas

TMMS	Transformer Monitoring/Management System
UK	United Kingdom

LIST OF SYMBOLS

%	Percent
kV	Kilo Volts
MVA	Mega Volt-Amp
MW	Mega Watts
°C	Degree Celsius
$\Re^n$	n-dimensional input space
n	Dimension of input space
m_i	Reference vector of neuron i
u_{ij}	j^{th} component of reference vector for neuron i
x	Input vector
ξ_k	k^{th} component of input vector
N_{ij}	j^{th} neighbourhood of neuron i
c	Best matching unit
t	Time-step
$h_{ci}(t)$	Neighbourhood kernel centred on c at t
$\alpha(t)$	Learning-rate function at t
$h(\|\|r_c - r_i\|\|, t)$	Neighbourhood function at t
r_c, r_i	Location vectors of neurons c, i
$\sigma(t)$	Width of neighbourhood kernel at t
num_neuron	Total number of neurons
dlen	Total number of input vectors
x-dim, *y-dim*	Map-sizes for x and y-dimension of SOM
α, β	Principal eigenvalues
$u(i)$	i^{th} component of u-matrix

CHAPTER 1

INTRODUCTION

1.1 Power Transformers and Their Reliability

Power transformers are among the most expensive and critical components of the electrical power system. Essentially, the power transformer converts a system of alternating voltage and current into another system of voltage and current of different values and at the same frequency for the purpose of transmitting electrical power throughout various points in the generation, transmission and distribution network.

The reliability of power transformers is of paramount importance to the profitable operation of power utilities, not least because the expenses of acquisition, transportation and installation of transformers are very substantial. In addition, their failures and subsequent unavailability may lead to loss of revenues and possible system constraint costs. These failures, when catastrophic, will further incur significant consequences in terms of peripheral equipment destruction, environmental damages and unplanned emergency utilisation of human resources and alternating power sources [1]. These are further reinforced by the fact that power utilities throughout the world are facing mounting challenges to supply uninterrupted power and efficient services to customers in today's increasingly competitive and deregulating energy market.

Fortunately, power transformers are reliable and durable items of electrical equipment. According to an international survey by CIGRÉ [2], typical failure rates

for large power transformers with windings for voltages up to 300 kV are in the range of 1% to 2% per annum. Nevertheless, power transformers are on middle ground of reliability when compared with other devices. Magnetic devices such as power transformers normally have a rate of 2.46 severe failures per 10^6 hours with insulation operating at 105 °C. This compares with a rate of 0.16 failures per 10^6 hours for composition resistors operating with rated power at 70 °C and a rate of 4 failures per 10^6 hours for paper capacitors operating with rated voltage at 85 °C [3].

Fundamentally, power transformers may fail in a variety of ways and for a variety of reasons. Generally, it is believed that power transformers conform to the "bathtub" failure-rate curve [3], as illustrated in Figure 1.1. It is characterised by a relatively high initial failure rate, sometimes referred to as infant mortality. This is due to design and manufacturing defects that tend to reveal themselves much earlier during operation. The failure rate decreases rapidly with time, approaching a low and constant value during most of the life of power transformers. This is the stress-related region in which transformers failed due to thermal, electrical and mechanical stresses in daily operation. As transformers approach the end-of-life (EOL), e.g. after 30 to 40 years, the failure rate rises and is attributed to the ageing of insulation materials within the power transformers. Some failure causes of power transformers that would initiate the foregoing failure modes are explained in Chapter 2.

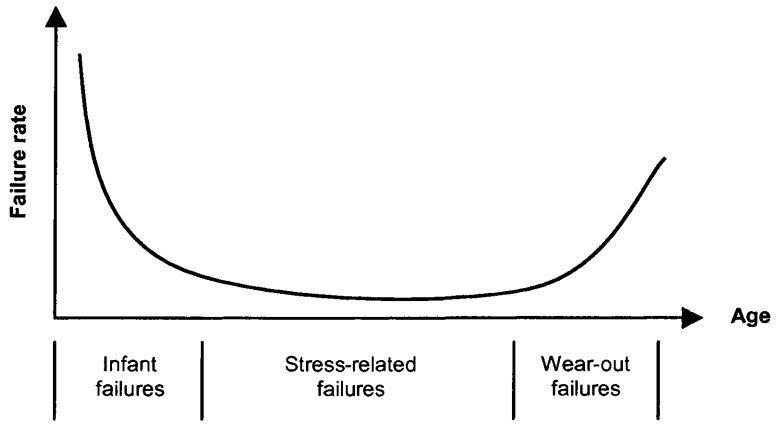


Figure 1.1: Failure rate versus age of the power transformers

1.2 Condition Assessment of Power Transformers

Power utilities throughout the world are increasingly focusing their attention on developments in condition assessment of power transformers, which includes fault diagnosis and condition monitoring on main units of power transformers and peripheral devices attached to them such as bushings and tap changers. The real driving forces behind all these interests are the increasing age of the transformer population and the deregulating and competitive environment of the energy market. Therefore, power utilities are prompted to economise on the operating and maintenance costs by utilising their assets fully and effectively. One of the means to attain substantial economic benefits is to switch from time-based maintenance (TBM) to condition-based maintenance (CBM) [4]. Nonetheless, the foregoing maintenance strategy requires the availability of reliable diagnostic methods and modern on-line monitoring approaches.

Diagnostic methods for power transformers have been well developed over the last 20 years. Improvements in analytical procedures and test equipments have allowed engineers to inspect the status of transformers in service by making routine and special checks on some key parameters that are associated with the health and conditions of power transformers. Diagnostic methods are developed as a result of theoretical, experimental and empirical knowledge about the transformer ageing and failure mechanisms.

Of all the established methods, dissolved gas analysis (DGA) is perhaps the best-known technique for diagnosing the incipient faults and monitoring the health of power transformers. This is accomplished by relating several ratios of key dissolved gases to some predefined conditions of power transformers according to several DGA interpretation schemes. Nevertheless, the DGA technique is not able to identify the location of faults once detected. Other diagnostic methods are capable of locating the faults, if there are any. Examples of such techniques are acoustic emission, infrared emission and radio frequency interference (RFI). The foregoing diagnostic

methods can be performed on-line, i.e. disconnection of transformers is not required. Apart from that, there is also a host of other methods being adopted by engineers that necessitate the disconnection of transformers from service, such as the frequency response analysis (FRA), power factor measurement, magnetising current measurement and polarisation spectrum. Both on-line and off-line diagnostic methods will be briefly discussed in the next chapter.

Recently, electronic sensors have been developed with an aim to continuously monitoring the operating conditions of power transformers; one can find commercial products to measure temperatures, gas-in-oil, moisture content, vibration and partial discharges (PD). These electronic equipments can be mounted onto the external part of power transformers, e.g. transformer tanks, with ease; sensor data can be downloaded to computers through direct cable-linkages or remote dial-ups via modems.

Major goals to be achieved by implementing the foregoing on-line monitoring approaches are to supervise the real-time operating condition and to detect the onset of abnormal circumstances or faults. There are various ways in which sensor measurements could be utilised; a classical way of analysing the sensor data is to define one or several thresholds for each measurement in order to set alarms. If a high-level alarm sets off, the operator concerned is alerted and diagnostic tests should then be performed so as to identify and locate faults. Alarm-based monitoring approaches are commonly used because they are non-intrusive and simple to install. However, the setting of thresholds is critical; they may be set too high and early fault detection may fail. On the other hand, when an alarm occurs too often (i.e. false alarm), its relevance might be altered. In case of a real fault, the transformer might be switched off too late and is subjected to damages with severe financial consequences. Thus, on-line monitoring approaches have to fulfil numerous requirements in order to justify the financial investments of power utilities. Basic pre-requisites are their suitability and reliability for real-time operation, reasonable installation and operating costs and good quality and stability of recorded measurements [5]. Various on-line monitoring approaches will be briefly discussed in the next chapter.

Artificial intelligence (AI) techniques have been employed extensively for enhancing the condition assessment practices of power transformers. In 2001, Kang et. al. [6, 7] reported the development of an automatic condition assessment technique for on-load tap changers (OLTCs) of power transformers based on the application of self-organising map (SOM). The SOM was used for classification and extraction of essential features from vibration signatures detected from an accelerometer that is mounted on the tap-changer tank. The foregoing technique has been developed into an on-line diagnostic system capable of indicating current status of condition and estimating the remaining lives of OLTCs.

In 2001, De et. al. [8] reported the application of SOM for extracting hidden features from frequency response patterns of windings of power transformers; a learning vector quantisation technique was also used to classify visually indistinguishable response patterns. This method has exhibited high diagnosis accuracy by successful detection and discrimination of impulse faults, of a different nature and site of occurrence, in windings of power transformers. In addition, the SOM was also applied for pattern recognition of fault patterns during the impulse testing of transformers, as reported by the same authors [9] in 2002. This technique can be applied for detection and diagnosis of the nature and location of fault in oil-filled power transformers during the impulse testing. Besides, the SOM can also be employed for short-term prediction of oil temperature change of an indoor transformer, as reported by Hong et. al. [10] in 2002.

Furthermore, a partial discharge (PD) localisation technique has been established by Werle et. al. [11] in 2001, which is based on neural-network (NN) classifiers trained by sectional winding transfer functions of transformer coils. This technique can be employed for determining apparent charges and location of the PD origin. Besides, the application of an adaptive neuro-fuzzy system has been reported by Roizman et. al. [12] in 1999, which was employed to predict the moisture content of solid insulation of power transformers based on on-line measurements of moisture characteristics of oil. Therefore, the insulation condition of power transformers can be monitored based on this technique. In 1998, Booth et. al. [13] provided an

assessment of generic capabilities of NNs, in both estimation and classification mode, for condition monitoring of electrical plant. Lastly, Wang et. al. [14] demonstrated an incipient fault diagnosis system based on the DGA data of transformers in 2000. The method utilised NNs and expert system (ES) for detecting and discriminating incipient faults of various nature and severity.

1.3 Challenges in Interpretation of Transformer Condition Assessment Data

As briefly discussed in the preceding section, power utilities are actively looking at diagnostic and on-line monitoring methods to assess the health and operating condition of power transformers. Even though equipments and procedures are well established for engineers to attain a reasonably accurate and reliable recording of data, there are further challenges with regard to confident interpretation of transformer condition based on this data. In essence, the data interpretation process is regarded as the most important aspect of any condition assessment practice since the prime motive of power utilities to invest substantially in these technologies is to reliably diagnose faults and to accurately monitor the condition of power transformers.

DGA is the most common diagnostic method used by various power utilities. Although well established procedures and equipments are in place for the extraction, quantification and recording of dissolved gases, not many progresses have been reported with regard to the interpretation of the DGA data of power transformers. Currently employed approaches are fundamentally based on ratios and concentrations of key dissolved gases, such as interpretation schemes introduced by Dörnenburg [15] and Rogers [16], and later by Duval [17, 18], IEC [19, 20] and CIGRÉ [21]. Although some of these interpretation schemes were since adapted by various power utilities in order to suit their own circumstances, basically the fundamental principles of these modified methods still do not deviate very much from foregoing interpretation schemes.

There are several challenges pertaining to the foregoing interpretation schemes. Firstly, uncertainty and ambiguity still exist as to which key-gas ratios should be calculated and on the credibility of the suggested limits of gas ratios; power utilities may have to decide and adjust these values heuristically and based on their own experiences until satisfactory diagnoses are achieved. Secondly, the heuristic and empirical nature of some of the foregoing interpretation schemes have brought about discrepancies in interpretation; application of different interpretation schemes on identical set of DGA data may produce diverse diagnosis of transformer condition, thereby causing confusion among power utilities. Thirdly, interpretation of transformer condition is sometimes impossible to achieve owing to the inability of some of these schemes to provide interpretation for every possible combination of ratio limits. Consequently, the interpretation of a DGA record may have to depend on expert judgement, which may instigate even more confusion since each expert may have his/her own idea on what is happening inside the transformer unit based on the dissolved-gas information. Finally, even if an interpretation of transformer condition is easily accomplished via the application of foregoing schemes, the information obtained is mainly text-based; there is no software tool for engineers to "visualise" the evolution of an incipient fault once detected. Therefore, further research into novel analysis and interpretation techniques is clearly needed in order to improve the confidence level in the diagnosis of transformer health and condition based on the dissolved-gas information.

Hitherto, the current practice of interpretation of sensor measurements, which are generated by the on-line monitoring approaches, is fundamentally based on the setting of threshold values. However, the problem is further compounded by the fact that each transformer is unique and a set of threshold limits cannot be equally applied to even identically configured transformers. Besides, there does not exist a single algorithm that could simultaneously analyse several sensor parameters and generate a summary with regard to the operating condition of power transformer. Therefore, it is cumbersome for engineers to visualise parallel time-sequences of several sensor parameters and subsequently deduce the operating condition of the power transformer. Consequently, a new approach that performs the foregoing monitoring

and summarisation task is clearly beneficial to those power utilities that have invested significantly in on-line monitoring technologies.

1.4 Objectives of the Project

This research project focuses on the investigation and development of an improved approach for the analysis and interpretation of the condition assessment data of power transformers with the ultimate aim of enhancing the condition assessment process. Main objectives of the project are:

- To perform detail investigation on conventional DGA interpretation schemes and new AI-based fault diagnosis approaches so as to ascertain the strengths and weaknesses of currently available solutions. These will be served as literature survey towards researching on an improved approach for the analysis and interpretation of the DGA data of power transformers.

- To perform a detailed investigation on the data mining (DM) approach and particularly on a DM method known as the SOM so as to ascertain its feasibility for an improved analysis and interpretation of the DGA data and on-line monitoring measurements. The investigation will focus on addressing some vital issues concerning the practical implementation of the SOM.

- To investigate the feasibility of the proposed approach for the analysis and interpretation of the DGA data of power transformers. Various aspects of the investigation will be carefully examined. Outcome of the investigation will be compared and validated against several well-established DGA interpretation schemes and some real fault-cases.

- To further apply the proposed approach for the analysis and interpretation of the DGA data of bushings of power transformers. The outcome of the

investigation will be analysed so as to ascertain the feasibility of the proposed approach for an improved condition assessment of transformer bushings.

- To extend the proposed approach based on the SOM for the analysis of sensor measurements as collected from an on-line monitoring system of the National Grid Company (NGC), UK. The study will focus on investigating the feasibility and effectiveness of the proposed approach for determining and summarising the hourly operating condition of a power transformer based on multiple sensor measurements.

1.5 Scope of the Thesis

Chapter 2

An overview of causes of failure and condition assessment practices for power transformers is presented in this chapter. Firstly, several common causes of transformer failures are presented. Secondly, some well-established diagnostic methods and recent on-line monitoring approaches are briefly described.

Chapter 3

This chapter presents a thorough review on various aspects of the DGA method. Firstly, some basic concepts of DGA are explained; these include introduction to the background theory and some related procedures for performing the DGA test. Secondly, a comprehensive literature survey on conventional DGA interpretation schemes and new AI-based fault diagnosis approaches, which are based on the DGA data of power transformers, is reported.

Chapter 4

This chapter discusses the implementation of several well-established DGA interpretation schemes for the DGA data of power transformers, such as Dörnenburg Ratios, Rogers Ratios, IEC Ratios, Duval Triangle and CIGRÉ Methods. In addition, comparative studies carried out on these schemes are reported with emphasis on their relative strengths and weaknesses.

Chapter 5

This chapter introduces a novel methodology known as the knowledge discovery in databases (KDD), in which DM is an important constituent of the KDD process. Specifically, this chapter focuses on a DM method known as the SOM algorithm. Various issues for the practical implementation of SOM are discussed and an illustrative example, which demonstrates the capability of the SOM, is presented. Besides, the suitability and potential benefits of the SOM for the analysis and interpretation of the condition assessment data of power transformers are also discussed.

Chapter 6

This chapter presents the outcome of some simple statistical analyses that were conducted on the DGA data of power transformers. The obtained DGA database courtesy of the NGC, UK was analysed using several simple statistical approaches. The aim of performing the foregoing analyses is to obtain a rough understanding on the statistical characteristics of the DGA data before it is subjected to a high-level analysis by the proposed approach.

Chapter 7

This chapter presents a detail investigation on the feasibility of the proposed approach for an improved analysis and interpretation of the DGA data of power transformers. Firstly, a preliminary study on some selected subsets of DGA data is reported. Secondly, a further analysis on the entire DGA database is presented. Based on these investigations, an improved interpretation of the DGA data in relation to the transformer health and condition is suggested, which was then validated by utilising several well-established DGA interpretation schemes and some actual fault-cases. Finally, the advantages of the proposed approach are presented and discussed.

Chapter 8

A study is reported in this chapter in which the proposed approach was further applied for the analysis and interpretation of the DGA data of transformer bushings. Firstly, the general design of a transformer bushing is illustrated. Secondly, a detail discussion on some challenges concerning the interpretation of the DGA data of transformer bushings is presented. Thirdly, the outcome of the analysis is presented, which was subsequently validated using a conventional approach and two actual fault-cases. Finally, the advantages of the proposed approach are presented and discussed.

Chapter 9

This chapter explores the potential of the proposed approach for the analysis of multiple sensor measurements as gathered from an on-line monitoring system of the NGC, UK. Essentially, similar procedures of analysis were applied and the outcome is reported and interpreted. A discussion is then presented that centred on the feasibility of the proposed approach for monitoring and summarising the hourly operating condition of power transformers.

Chapter 10

The work conducted in this research project is summarised and concluded in this chapter, which emphases on some major achievements and progress that have been attained. In addition, possible future work with regard to the current research is suggested.

1.6 Contributions from this Research Project

The main contributions from this research project are stated below:

- A comparative study of existing dissolved gas analysis (DGA) methods and their codification for transformer gas ratio analysis.

- The production of a structured method for the use of data mining (DM) methods in general, and self-organising maps (SOMs) in particular, in transformer condition monitoring data classification applications.

- The development of a robust analytical tool-set capable of processing large, heterogeneous transformer data sets and specific transformer type monitoring data in order to identify incipient fault developments.

- A detailed comparative analysis of conventional DGA interpretation methods with the proposed SOM technique.

- A visualisation approach used to track transformer condition and fault evolution.

- All of the proposed methods have been validated and tested on actual data supplied by the National Grid Company (NGC), UK.

CHAPTER 2

CAUSES OF FAILURE AND CONDITION ASSESSMENT OF POWER TRANSFORMERS

2.1 Introduction

Power transformers are critical components of the electrical power system. Their failures and subsequent unavailability can cause significant consequences both financially and environmentally. Hence, it is important to understand the causes of these failures. In addition, these failures can be prevented effectively via the adoption of condition assessment practices such as diagnostic methods and on-line monitoring approaches. An overview of causes of failure and condition assessment practices is presented in this chapter.

2.2 Causes of Failure of Power Transformers

While the power transformer is undoubtedly one item of electrical equipment that is least subjected to breakdown when compared with other electrical devices within the substation, failures do occur from time to time due to various reasons. By definition of the IEC, failure means the termination of the ability of an item to perform a required function [20]. In case of power transformers, failures that most commonly arise in practice can be broadly classified into three categories, i.e. infant failures, stress-related failures and wear-out failures. Some common causes for these failures are described in following sections.

2.2.1 Infant Failures

Infant failures can be initiated by deficiency in either design or manufacturing process. It has been known that a common reason for early failure is the lack of compatibility between the design and usage of the transformer. This can be due to the failure of the designer and user to communicate effectively or as a result of some unforeseen stresses that the transformer is incapable of withstanding. Furthermore, the designer may fail to provide for specified or implied condition that the designer would normally be expected to consider. Thus, when reliability problems develop in a new design, the relationship between the design and usage must be investigated. Although neither the designer nor the user may want to accept responsibility, if the problem is to be resolved then the facts must be established since it should not be very surprising that reliability problems do arise in newly installed transformer. Apart from the design process, improper manufacturing and handling of components such as windings and oil-tanks during the manufacturing process could also lead to failures of power transformers.

2.2.2 Stress-Related Failures

Power transformers may fail due to electrical, thermal or mechanical stresses in normal operation. While the failure-rate associated with stress-related failures is relatively low and constant during most of the life of power transformers, they should not be ignored because the causes of such failures can actually be detected well before any failure occurs. Some of the common causes are listed below.

2.2.2.1 Discharge

Discharges within the power transformer can occur between conductors at different potentials, resulting in the destruction of insulation and electrical failure. It is often described as arcing, breakdown or short circuits. The potential can be induced by the transformer or it can be externally applied, for example, voltage spikes caused by lightning strikes. Induced voltage can lead to discharges between portions of the

winding in which the voltage is induced, or it can cause discharges from a portion of the winding to ground or other windings. Externally applied voltage other than induced voltage can also cause discharges between the windings to which the voltage is applied and other windings or ground. Discharge is usually accompanied by a high current, causing overheating and carbonisation along the passage, permanently reducing the electrical resistance. Furthermore, openings can develop in windings in the discharge region owing to the concentrated heat. Discharge is a highly localised condition and the carbon path can usually be seen if the device is carefully dissected. Identification of the discharge passage is important for determining corrective action.

2.2.2.2 Partial discharge

Partial discharge (PD) is a phenomenon that only partially bridges the insulation between conductors. Specifically, corona is referred to as a form of PD that usually occurs in the gaseous medium around conductors that are away from solid or liquid insulation. The ionisation associated with PD is insufficient to establish a low resistance path between conductors allowing a high current to flow as in the foregoing disruptive discharges. A power transformer will still function with the presence of PD. In fact, the PD may go undetected until a failure occurs. PD is not usually a problem until the potential between conductors reaches several thousand volts. It is sometimes a problem in sensitive devices because it raises the noise level. Approaches have been developed that detect PD. As qualitative determiners they work well, but they do not provide much useful information for determining acceptable limits. The problem is very complex. The point where PD occurs and the type of insulation presents will alter the effect of PD on insulation life.

2.2.2.3 Thermal fault

Thermal fault (TF) is a general term for overheating and hotspots. It is due to excessive temperature rise in insulation, which leads to a decrease in resistance and hence allows the circulation of leakage currents that further increases the insulation temperature. Insulation damage develops, culminating in a dielectric breakdown. TF

is a diffused condition in which internal examination of overheated coils will show discolouration over a large region. Discharges will occur in this region at a point that subjected to particularly high electrical stress. Careful analysis is sometimes necessary to determine whether the fault is a discharge or a TF when specimens show evidence of both problems. TF will accelerate the ageing rate of the insulation.

2.2.2.4 Open windings

Openings develop in windings wound with very fine wire. This problem is distinct from openings resulting from the melting of conductors associated with discharges. Winding conductors are fragile and becomes increasingly so during handling because of the tendency for copper to work-harden. The assembly processes of winding and lead dressing severely stress the wire. Subsequent thermal changes and vibration can fracture the wire previously weakened by handling during manufacture, a common cause of failure.

2.2.2.5 Overstressing

Power transformers are frequently subjected to overstress conditions. The effect of overstressing is not determinable. Common examples of overstressing are overloading or shorting of secondary windings (pending the operation of primary circuit protection), electrical overstressing as an acceptance test procedure, and high-voltage transients occurring during operation. These overstress conditions do not result in immediate failure, but the damage they cause may result in subsequent failure during normal operations. This situation is difficult to deal with because the complete operating history of the devices is seldom known. Failure analysis and corrective action may have to rely on statistical data.

2.2.2.6 Shorted turns

Electrical stress exists between turns as a result of induced voltages. The only insulation between adjacent turns is the insulating film on the magnet wire. Normally

the voltage between adjacent turns is low, and the insulating film is adequate to withstand the voltage. Shorts develop under temporary over-voltage condition in regions where insulation is weak. A particularly serious condition is a crossover in which a turn in the winding crosses over a previous turn instead of remaining in an even helix. The area of contact between turns in a crossover is subjected to excessive pressure and electrical stress. Shorted turns can also occur between adjacent layers of layer-wound coils by puncture of layer insulation. The electrical stress between layers exceeds the stress between adjacent turns. Under steady state condition the induced voltage across a winding distributes itself uniformly across the turns of the winding. Under transient condition, the voltage will tend to develop across the end turns of the winding, resulting in an electrical stress many times greater than normal.

2.2.3 Wear-Out Failures

Wear-out failures in power transformers occur as a result of the combination of ageing effects of time, temperatures, and electrical stresses on insulation. Additional factors contributing to wear-out are mechanical stresses from repeated thermal cycling and vibration of conductors due to the varying currents they carry and their presence in the magnetic field.

2.3 The Need for Condition Assessment

Failures of power transformers are problematic for two reasons. Firstly, transformer failures lead to operational and financial difficulties owing to the interconnected nature of a power system and the significant cost of repairing or replacing the failed transformers. Secondly, power transformers are encased in tanks of flammable and environmentally hazardous fluid; catastrophic failures are often accompanied by fire and spillage of this fluid, thus presenting high risks to human, other peripheral equipments, surrounding properties and the local environment. Therefore, there are clear incentives for power utilities to embrace condition assessment practices for their transformer population.

Generally, the practice of condition assessment on power transformers include fault diagnosis and condition monitoring on main units of power transformers and peripheral devices attached to them such as bushings and tap changers. These practices are advantageous based on several grounds. Firstly, the adoption of condition assessment enables better and more efficient utilisation of resources. Some failures in power transformers are not detectable owing to their sudden nature. However, some slowly evolving faults, which are generally referred to as incipient faults, can be detected and diagnosed such that remedial actions can be taken at the most commercially opportune moment [22]. In fact, it is possible to continue operating the power transformer even when an incipient fault has been detected since the condition assessment practice will provide a means of tracking vital parameters of the transformer so that further abnormal or problematic behaviour can be detected and corrective action taken.

Secondly, condition assessment also allows the prediction of remaining useful life of power transformers. Data from the assessment process can be associated with actual condition of various vital components inside power transformers such as windings and insulation. Finally, condition assessment allows the effective deployment of condition-based maintenance (CBM), in which greater financial savings can be expected if compared to time-based maintenance (TBM).

2.4 Diagnostic Methods

Generally, diagnostic methods are applied for fault diagnosis of power transformers. However, some of these methods can also be used for condition monitoring purposes, such as the dissolved gas analysis (DGA) and furfural analysis. These methods are developed as a result of theoretical, experimental and heuristic knowledge about the transformer ageing and failure mechanisms.

Diagnostic methods can be broadly classified into two categories: on-line and off-line approaches. The on-line diagnostic approaches do not require the disconnection

of transformers from service while the off-line approaches necessitate the disconnection of transformers and are generally more costly to implement. Besides, these methods can also be identified as either "routine tests" or "special tests" [23]. "Routine tests" are normally carried out on all transformers on a periodic basis for screening purposes in order to detect incipient faults and to ascertain general condition. In contrast, "special tests" are applied only as required for fault diagnosis and detailed investigation in response to one of a set of triggering circumstances which include the following:

- To investigate a poor routine test result.
- Following a protection operation indicating an internal fault.
- Following a system event that might have caused damage.
- As part of a commercial asset review.
- To decide whether a redundant asset is worth retaining.
- Before and after oil reclamation in order to determine the effectiveness of the treatment.
- Before scrapping so as to correlate test results with the observed condition.
- To establish the results as expected from a "normal" transformer unit.

Some of the common diagnostic methods are listed in Table 2.1 [23] and are briefly explained in following sections.

2.4.1 Dissolved Gas Analysis

DGA is one of the most common diagnostic tools adopted by power utilities. This approach is based on the mechanism of formation of gases from the degradation of mineral oil and cellulose insulation during the onset and evolution of faults, which are signified by the increase in either localised or diffused temperatures. IEC recommends the extraction and quantification of nine dissolved gases from the oil sample, and these are: oxygen (O_2), nitrogen (N_2), carbon dioxide (CO_2), carbon monoxide (CO), hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4) and

Table 2.1: Diagnostic methods for power transformers

Test	Diagnosis	Location of Fault	Cost and Convenience	Usage
Dissolved gas analysis	Discharges Overheating	None, integrates over time	Cheap and easy on-line	Routine test
Furfural analysis	Paper ageing Hotspots	None, integrates over time	Cheap and easy on-line	Routine test
Radio frequency interference	Discharges	None or partial	Cheap and easy on-line	Routine test
Acoustic emission	Discharges	Good if discharges not in winding	Moderate cost on-line	Special test
Infrared emission	Tank currents Cooler blockages	Good	Moderate cost on-line	Routine test Special test
Frequency response analysis	Mechanical condition	Good, can identify phase and winding of fault	Expensive, disconnection required	Special test
Polarisation spectrum	Paper moisture/ageing	None, although may indicate presence of wetspots	Expensive, limited disconnection required	Special test Routine test at major maintenance
Power factor	Oil/paper insulation condition	Partial, depends on how windings can be separated for test	Expensive, disconnection required	Special test
Magnetising currents	Shorted turns and some core faults	Partial, can indicate phase of faults	Expensive, disconnection required	Special tests
Turns ratio	Shorted turns	Partial, can indicate phase and winding of fault	Expensive, disconnection required	Special test
Winding resistance	Winding joint and tap-changer selector contact problems	Partial	Expensive, disconnection required	Special test

acetylene (C_2H_2) [24]. Two types of incipient faults are detectable: electrical fault (EF) (such as discharges and PD), and TF (such as hotspots and overheating). The association between dissolved gases and the foregoing incipient faults is established by referring to several well-established DGA interpretation schemes, such as the IEC Ratios [19, 20], Rogers Ratios [16], Dörnenburg Ratios [15], CIGRÉ Methods [21] and the Duval Triangle [17, 18]. The interpretation is normally accomplished via the computation of several ratios of key dissolved-gases. An overview of the DGA approach will be presented in Chapter 3.

2.4.2 Furfural Analysis

Furfural analysis by high performance liquid chromatography (HPLC) has gained increasing preference as a means of estimating the degradation process of paper insulation in power transformers. Generally, degradation of cellulose insulation such as paper will produce a range of furans, which will dissolve in mineral oil. The advantage of using furans instead of concentration of CO_2 and CO is that the former are specific to paper degradation only; degradation of oil will not produce furans. The concentration and rate of change of furanic compounds indicate the condition and the rate of ageing of paper insulation. In addition, the benefit of using furfural analysis is that it is non-intrusive, and hence disconnection of power transformers are not required as the concentration of furans is simply determined through HPLC on oil samples.

2.4.3 Radio Frequency Interferences

The electromagnetic radiation generated by internal discharges occurring in transformers can be measured via the radio frequency interferences (RFI) technique. This can be accomplished by using a radio frequency (RF) spectrum analyser that is equipped with different kinds of antennas [25]. A computer is then used to control measurements and to store the measured data. The RFI technique is a useful tool for detecting discharges and is especially advantageous for measurements on-site since it is not necessary for the disconnection of power transformers.

2.4.4 Acoustic Emission

Acoustic emission refers to the generation of transient waves during the rapid release of energy from localised sources within a material. The sources of acoustic emission in power transformers are mainly due to discharges and PD. This technique is based on the detection and conversion of these high frequency waves into electrical signals. This is accomplished by directly coupling the piezoelectronic transducers onto the surface of the structure under test. The output of each sensor is then amplified through a low-noise pre-amplifier, filtered to remove any extraneous noise and further processed by using suitable electronic equipment. If more than one sensor is used, the acoustic emission source and thus the fault location can be located through the application of triangulation method. Advantages of the acoustic technique are high sensitivity, rapid detection of faults and minor disturbances in power transformers on-site.

2.4.5 Infrared Emission

The infrared emission technique is applied for locating abnormal heating within the power transformers. It is accomplished through the use of thermal imaging cameras. Essentially, a thermal scan on transformer main-tank and associated peripheral equipments such as load tap-changer (LTC) and bushings can reveal LTC heating problems, radiator blockage, hotspots and overheating in tanks. The infrared emission technique will detect and translate the heat emission into electronic signals, which can then be imaged, measured and analysed. The inspection of heating problems using this technique is fast and can be performed on-site easily without the disconnection of power transformers.

2.4.6 Frequency Response Analysis

During its lifetime, a transformer can be subjected to several short circuits, either internal or external, with high fault currents. The forces of these short-circuit currents may cause deformation or displacement of the winding assemblies. As deformation

results in minor changes to the internal inductance and capacitance of the winding structure, a change in the characteristic frequency response can be detected at the terminals of the transformer by frequency response analysis (FRA) [4]. It can be accomplished by using a commercially available network analyser. The frequency response of a transformer can be analysed off-line by connecting cables from the analyser to one terminal of the winding and the other cable connecting from ground via current transformer (CT) to another terminal of the winding. Winding deformation can be detected by analysis and comparison of frequency response characteristics. The FRA method has been proven to be sensitive in detecting typical winding faults and it is immune to electromagnetic interference.

2.4.7 Polarisation Spectrum

The presence of moisture in the insulation system will accelerate the ageing rate of the insulation. Since water molecules are polar in nature, a test method has been developed to assess the relationship between ageing, moisture content and polarisation response [26]. Essentially, an insulation system is polarised by applying a DC voltage. When the terminals are subsequently short-circuited, a residual polarisation may remain when the short-circuit is removed. This residual polarisation results in a recovery voltage. The interpolated curve formed from the measured recovery voltages versus charging period is known as the polarisation spectrum. Since the recovery voltage will change according to the insulation condition, there is a good correspondence between measured characteristics and the state of ageing or the moisture content of the insulation. The condition of a particular insulation system can be characterised by comparing its polarisation spectrum with other insulation systems of different quality.

2.4.8 Power Factor

The power factor of the insulation is the measure of its dielectric loss. This diagnostic method can be used for indicating the overall condition of the insulation, i.e. the degree of degradation. The power factor method is particularly suitable for

detecting the presence of moisture and other contaminants inside the transformer windings. It is well known that the presence of these materials will reduce the dielectric strength of insulation, making them more susceptible to electrical stresses.

2.4.9 Magnetising Currents

The magnetising current in a normal transformer is about ten percent of the full load current. During internal winding faults, depending on the location of the fault, the magnetising current increases rapidly. The distribution of transformer current subsequent to an internal electrical fault thus differs totally from the distribution of normal load or no-load currents and is governed mainly by the internal reactance of the windings. Therefore, winding faults such as shorted turns can be identified through the measurement of magnetising currents.

2.4.10 Turns Ratio

Test is conducted on power transformers to ensure the turns-ratio of windings is correct; shorted turns may have occurred if an incorrect turns-ratio is obtained. This test is performed using a ratiometer, which itself consists of a single-phase, double-wound transformer, having a constant primary winding and a variable secondary winding [27]. The turns-ratio is measured via a special configuration of connection between windings of the test transformer and windings of the transformer in the ratiometer.

2.4.11 Winding Resistance

The DC resistances of both HV and LV windings can be measured simply by the voltmeter/ammeter method. This information provides the data necessary to permit the separation of power and eddy-current losses in the windings. This is necessary in order that transformer performance may be calculated at any specified temperature.

2.5 On-Line Monitoring Approaches

On-line monitoring approaches are used for real-time supervision of power transformers. As distinct from diagnostic methods, these on-line monitoring approaches are only capable of detecting the onset of faults or abnormal circumstances within power transformers; diagnostic methods must be used to identify the nature and location of faults once they are detected. Essentially, analogue and electronic sensors are fitted onto transformers; recorded data can be transmitted to computers and software packages are available to present the recorded information to users. There are two motives for implementing on-line monitoring approaches, i.e. to detect the onset of faults through the monitoring of important parameters, and to allow the effective deployment of condition-based maintenance (CBM).

In essence, key parameters have to be determined which can effectively characterise and reflect the various stresses that could affect the rate of ageing and the operating condition of power transformers. Nonetheless, the selection of these parameters must be based on failure statistics of power utilities and expected consequences of failures [28]. Table 2.2 illustrates a typical failure distribution for large substation transformers equipped with on-load tap changers (OLTCs) [2]. As observed from the table, it can be concluded that vital parts to be monitored are windings, which include the oil/paper insulation, and OLTCs.

Table 2.2: Typical failure distribution for substation transformers with OLTCs

Part	Failure Distribution
OLTC	41%
Windings	19%
Core	3%
Terminals	12%
Tank/Fluid	13%
Accessories	12%

In essence, key parameters that are related to windings and the insulation system are gas-in-oil, moisture content and PD. As for the OLTCs, several alternatives are possible; however, vibration monitoring seems to be the most suitable candidate [28]. In addition, monitoring of the top and bottom oil temperatures and the winding temperatures are regarded as base information and should also be included as part of the key parameters. Several on-line monitoring approaches that offer the measurement and monitoring of foregoing parameters are briefly explained in following sections.

2.5.1 Gas-in-Oil Monitors

Generally, two developments can be distinguished for the on-line gas-in-oil monitors: sensors for monitoring mainly hydrogen as the most important gas generated during onset of faults, and sensors that are capable of monitoring several gases simultaneously [5]. Basically, the hydrogen sensor allows only a basic trend analysis, in which thresholds can be set for various alarm levels. One such monitor is the Hydran 201R developed by GE-Syprotec [29], in which fuel-cell technology is used. Other more advanced sensors are capable of monitoring several gases simultaneously, which open up the possibility for on-line diagnosis. Examples of such development are presented by ABB (metal-oxide technology) and by Serveron (semiconductor technology).

2.5.2 Measurement of Moisture Content

A continuous measurement of relative saturation of moisture in oil can provide valuable information to a transformer engineer. For instance, a transformer that is heavily loaded during the day, but which cools down rapidly at night, will exhibit very high relative saturation of moisture in oil. Instruments to measure moisture content of transformer oil are made by Doble Engineering and Panametrics.

2.5.3 Measurement of Partial Discharges

Acoustic sensors have been used for monitoring the occurrence of PD. The benefit of using acoustic monitoring is its ability to identify the PD location. This is accomplished by fitting sensors at many points around the tank; signals received from the sensors can be used for locating the fault through the triangulation process [29].

2.5.4 Measurement of Temperatures

The top and bottom oil temperatures can be measured on-line using currently available sensors. Although this method provides no direct information about the hotspot temperature of windings, it can be calculated according to the IEC guide on the loading of power transformers. A more direct measurement of hotspots can be accomplished via the fibre optic sensors, which can be installed in new transformers. In general, two main types of sensors are used: fibres that measure the temperature at one point, and distributed fibres that measure the temperature along the length of the winding.

2.5.5 On-line Monitoring of Bushings and OLTCs

Transformer bushings consist of multiple alternating layers of foil and oil-impregnated paper. A very small charging current flows continuously when the system is energised. Changes in the charging current are an indication of the degradation of the insulation system. The charging current can be monitored on-line by a device that plugs into an existing capacitor tap of the bushing. An example of such device is the InsAlert monitoring probe by Square D Co. [29]. On the other hand, vibration sensors can be used for monitoring of OLTCs. It has been shown that both mechanical and electrical faults can be detected, as well as wear of contacts and changes in transition times.

2.5.6 On-line Monitoring System: An Integrated Package

On-line monitoring systems are now available for continuously monitoring various vital parameters of power transformers in a neat and unified package. Examples of such systems are the ABB's T-Monitor [30] and the GE-Syprotec's Transformer Monitoring/Management System (TMMS) [1]. Generally, an on-line monitoring system comprises of three main elements: sensors, analogue-to-digital conversion of the measuring quantities and a computer for data processing, analysis and presentation. Transformer parameters that can be monitored include currents, voltages, tap-changer position, temperatures, gas-in-oil, moisture content, operating condition of pumps and fans, velocity of oil flow and oil levels. Apart from direct measurement and presentation of parameters, these on-line monitoring systems are also capable of determining the operating condition of transformers through the use of models [1]. A simplified diagram of an on-line monitoring system is illustrated in Figure 2.1 [1].

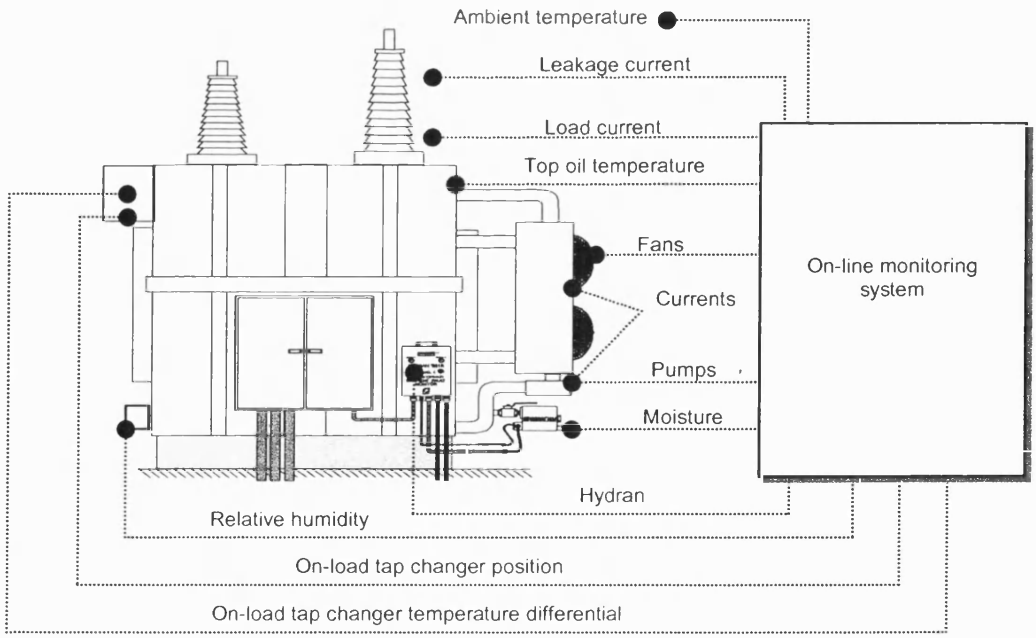


Figure 2.1: A simplified diagram of an on-line monitoring system

2.6 Summary

An overview of common causes of failures and condition assessment approaches has been presented in this chapter. Essentially, power transformers may fail under three common causes: infant mortality, operating stresses and ageing. Diagnostic methods are mainly used for fault diagnosis of power transformers, i.e. to identify the type and/or location of faults. In contrast, on-line monitoring approaches are used for real-time supervision of power transformers, in which abnormal circumstances or faults detected must be followed up by diagnostic tests. Nonetheless, effective deployment of the foregoing condition assessment practices can prevent transformer failures and hence present significant financial savings for power utilities.

CHAPTER 3

DISSOLVED GAS ANALYSIS

3.1 Introduction

Various aspects of the dissolved gas analysis (DGA) are presented in this chapter, which comprise of the related theory, procedures of the DGA test and a literature survey on conventional DGA interpretation schemes and new artificial intelligence (AI) based fault diagnosis approaches that are based on the DGA data of power transformers. These foregoing aspects will serve as an introduction to DGA and to the recent developments of its application for fault diagnosis of power transformers.

3.2 Background Theory

Insulation oil of power transformers is a mixture of many different hydrocarbon molecules containing CH_3, CH_2 and CH chemical groups, which are linked together by carbon-carbon molecular bonds. The onset of either an electrical or a thermal fault will cause the scission of some of the carbon-carbon and carbon-hydrogen bonds with the formation of small unstable fragments. These fragments, either in radical or ionic form, will recombine rapidly through complex reactions into gases. When the fault concerned is not severe, the gases formed will dissolve in oil, with a small proportion diffusing from the oil into any gas phase above it. In case of a severe fault, the generation of gases is so rapid that they rise to the gas phase immediately without dissolving much in the oil. Gas relay will be activated in the latter case to disconnect the faulted transformer in order to prevent it from further damage.

In 1970, Halstead performed a theoretical thermodynamic assessment of the formation of hydrocarbon gases in the insulation oil [31]. It was suggested that on the basis of equilibrium pressures at various temperatures, the proportion of each hydrocarbon gas in comparison with each of the other hydrocarbon gases varied with the temperature at the point of degradation. This has led to the assumption that the rate of evolution of any particular hydrocarbon gas varied with temperature, and that at a particular temperature there would be a maximum rate of evolution of that gas, and that each gas would attain its maximum rate at a different temperature. Study of the Halstead thermodynamic equilibria has suggested that with increasing temperature, the maxima would be in turn methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4) and acetylene (C_2H_2). In addition, the rate of evolution of hydrogen (H_2) will continue to rise with temperature. Figure 3.1 illustrates a simplified representation of this hypothesis [16].

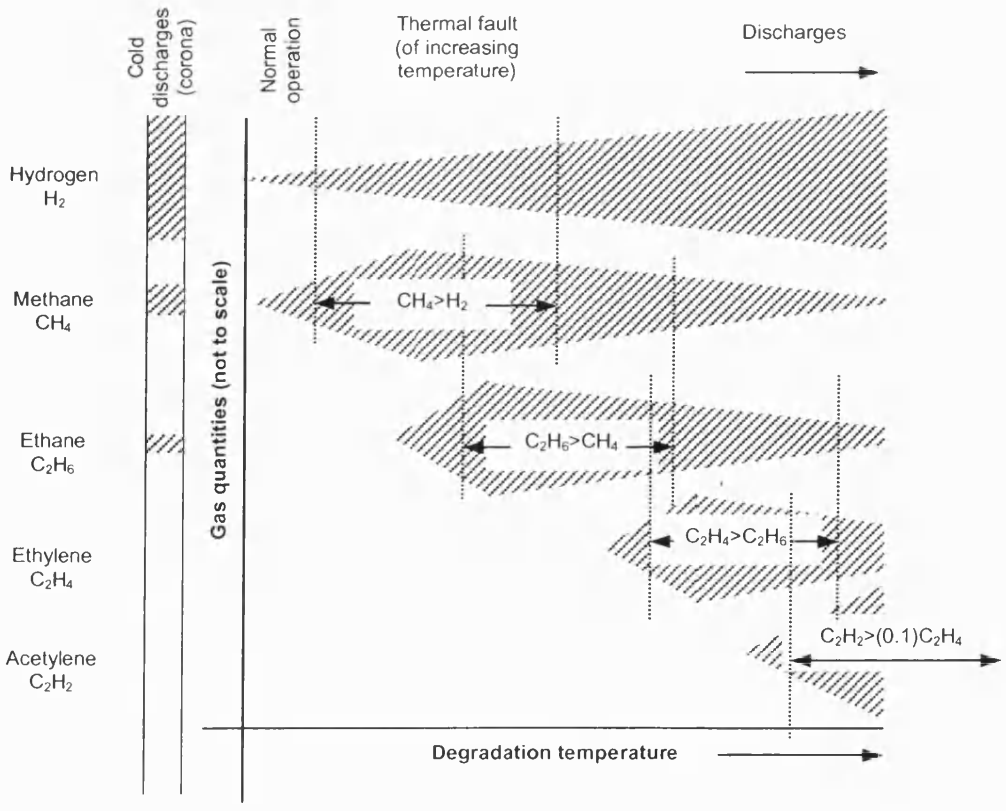


Figure 3.1: Comparative rates of gases evolution from oil

Therefore, close correlation can be established between the relative composition of dissolved gases and the type and severity of fault. Generally, two distinct incipient faults can be identified through DGA: electrical fault (EF) and thermal fault (TF); EF can be further categorised into discharges and partial discharges (PD) and TF is often referred to as overheating or hotspots.

Generally, the onset of low-energy fault such as PD of the cold plasma type (i.e. corona) will produce H_2 as the major dissolved gas. In other cases, degradation of insulation oil is mainly caused by heat, with variation in the type of hydrocarbon gases produced as temperature rises. A little degradation of oil at normal operating temperature produces mainly H_2 and CH_4. Higher temperatures and higher energies are caused by TF such as overheating or hotspots. Thus, temperature from little above normal operating temperature (say, 150 °C) will produce mainly CH_4. As fault temperature rises, both C_2H_6 and C_2H_4 will be produced in increasing quantities. In case much higher temperatures (say, 800 °C to 1200 °C), occurring at discharges such as flashover or sparking, the production of C_2H_2 becomes significant. Note that the production of H_2 will also increase continuously with fault temperatures. If the foregoing faults also involve cellulose insulation such as paper, large quantities of carbon monoxide (CO) and carbon dioxide (CO_2) will also be generated.

Nevertheless, although the formation of some gases is favoured depending on the foregoing descriptions, mixtures of dissolved gases are always obtained in practice. However, the ensued gaseous composition essentially matches the foregoing descriptions on mechanisms of gaseous formation in insulation oil.

3.3 Procedures of DGA Test

Detailed procedures for performing the DGA test are outlined in the IEC 567 Standard [24]. Generally, a DGA test comprises of three stages: sampling of oil, extraction of dissolved gases from oil and identification of dissolved gases. These procedures are briefly described as follows.

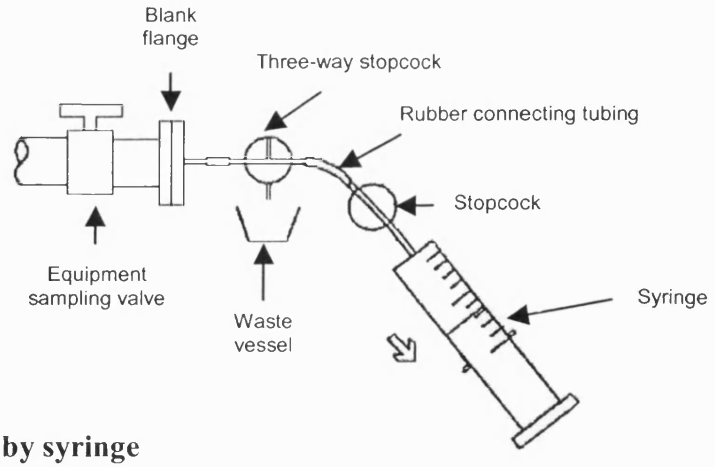
3.3.1 Sampling of Oil

The DGA test on power transformers requires sampling of a small volume of oil from transformer tanks. The sampling of oil can be done by using syringe, glass or stainless steel tube, and glass bottles. The choice of a suitable method depends on the apparatus available and on the quantity of oil needed for analysis. Oil sampling by syringe, as illustrated in Figure 3.2 (a), is the most convenient method and is suitable irrespective of the mode of transport of samples. On the other hand, sampling into a glass or stainless steel tube is also suitable, as illustrated in Figure 3.2 (b). However, the glass tube must be fitted with sufficient length of rubber tubing to act as an expansion chamber. Lastly, the sampling can performed using the glass bottle, as illustrated in Figure 3.2 (c), provided that the bottle is fitted with a suitable cap that allows oil expansion. Oil sampling using glass bottle is adequate for many purposes such as routine sampling on a large-scale basis from power transformers on-site.

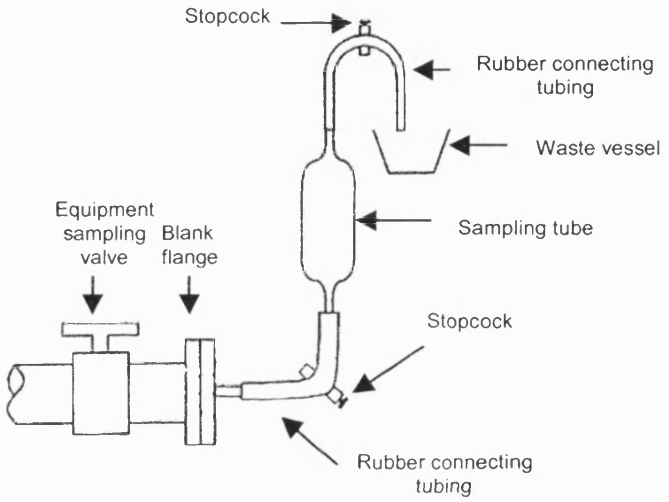
The selection of points at which oil samples are drawn should be decided with care. Normally, oil samples should be extracted from a point where it is representative of the bulk of oil in the transformer, and this is normally located at the bottom of the transformer tank. However, it will sometimes be necessary to draw samples deliberately where they are not expected to be representative in order to identify the location of a suspected fault.

3.3.2 Extraction of Dissolved Gases from Oil

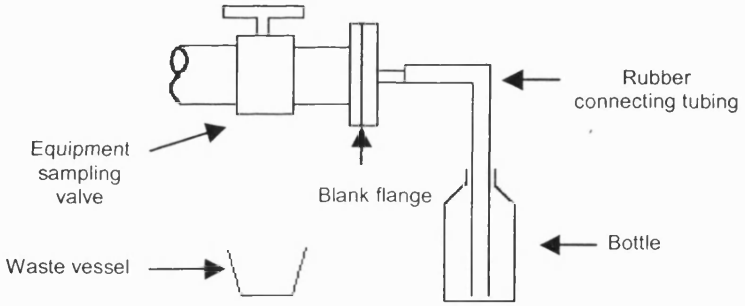
The extraction of dissolved gases is performed in the oil laboratory. Removal of dissolved gases can be done either by stripping or by vacuum extraction, as illustrated in Figures 3.3 (a) and 3.3 (b). The vacuum extraction method can be further identified into that of multi-cycle and single-cycle. The former method can remove about 97% and above of the dissolved gases and the latter about 90% to 99.8%. Conversely, the stripping technique is the simplest extraction method, where a carrier gas such as helium is bubbled directly through the oil.



(a) Sampling by syringe

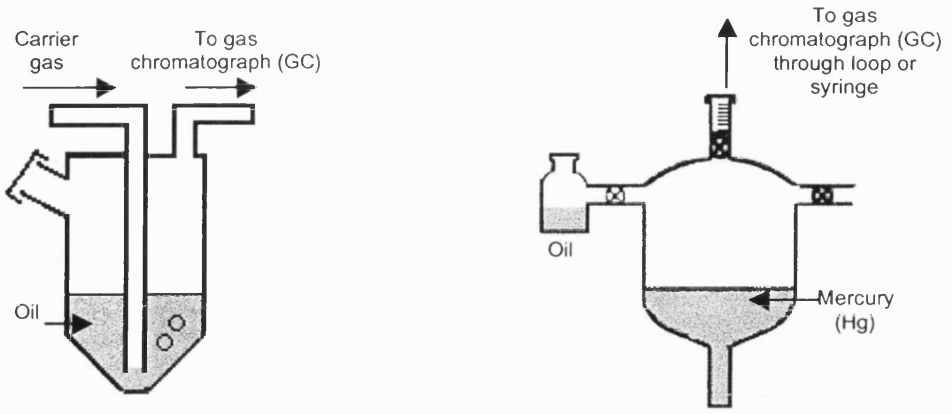


(b) Sampling by glass/steel tube



(c) Sampling by glass bottle

Figure 3.2: Oil sampling methods



(a) Extraction by stripping **(b) Extraction by vacuum extraction**

Figure 3.3: Dissolved-gas extraction methods

3.3.3 Identification of Dissolved Gases

The dissolved gases, after being removed from transformer oil via the foregoing extraction methods, are transferred to gas chromatograph (GC) for identification and quantification. According to the IEC 567 Standard [24], gases that need to be determined are: H_2, O_2, N_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO and CO_2.

A GC normally comprises of several components, as illustrated in Figure 3.4. The identification process is started with two separate runs; the extracted dissolved gases are first passed through the Porapak column and in the second run are passed through the molecular sieve column. Since a single detector having acceptable accuracy for detecting all gases is not available, gases eluted from foregoing columns are channelled through a thermal conductivity detector and a flame ionisation detector. The thermal conductivity detector can detect N_2, O_2, CO, CO_2 and H_2 while the flame ionisation detector can detect the presence of hydrocarbons such as CH_4, C_2H_6, C_2H_4 and C_2H_2.

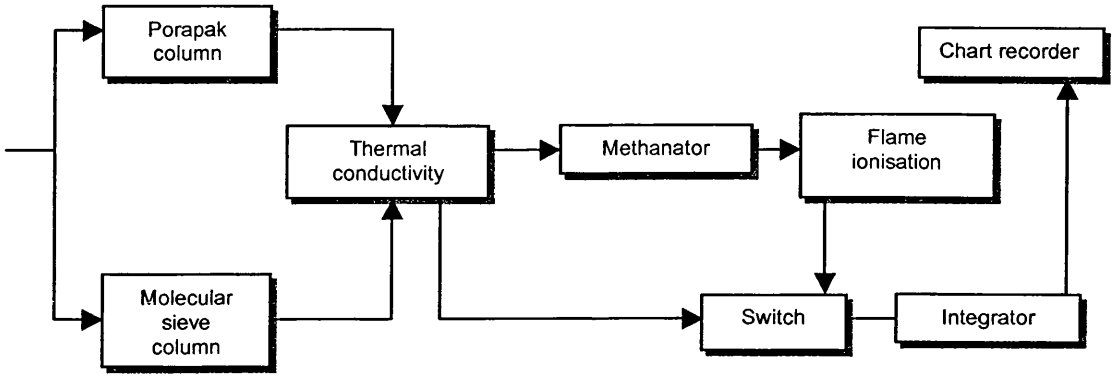


Figure 3.4: Simplified diagram of a gas chromatograph (GC)

Moreover, in order to improve the sensitivity towards CO and CO_2, a methanator is fitted at the inlet of the flame ionisation detector to convert CO and CO_2 to methane, which is then detected by the flame ionisation detector. The electrical output from both detectors is fed to an integrator with a chart recorder. The type and quantity of dissolved gases are recorded in part-per-million (PPM), in which 1 PPM means one micro-litre (μl) of a particular dissolved gas in 1 litre (l) of oil.

3.4 DGA Interpretation Schemes

The DGA data of power transformers as gathered from DGA tests is submitted for interpretation by oil experts or transformer engineers, who will infer the condition of a particular power transformer from the composition of dissolved gases based on several DGA interpretation schemes. Amongst these schemes, five most well-known and established methods are: Dörnenburg Ratios [15], Rogers Ratios [16], IEC Ratios [19, 20], Duval Triangle [17, 18] and CIGRÉ Methods [21].

Some other DGA interpretation schemes also exist that are utilised by power utilities in various countries. However, fundamental principles of these schemes are very similar to the interpretation schemes as mentioned above.

Before elaborating on foregoing DGA interpretation schemes, it is best to clarify several commonly encountered terminologies in these schemes. Definitions as listed in Table 3.1 are extracted from the IEC 60599 Standard [20]:

Table 3.1: Definitions for commonly encountered terminologies in DGA interpretation schemes

Term	Definition
Fault	An unplanned occurrence or defect in an item that may result in one or more failures of the item itself or of other associated equipment.
Non-damage fault	A fault that does not involve repair or replacement action at the point of the fault.
Damage fault	A fault that involves repair or replacement action at the point of the fault.
Failure	The termination of the ability of an item to perform a required function.
Electrical fault	A partial or disruptive discharge through the insulation
Partial discharge	A discharge that only partially bridges the insulation between conductors. It may occur inside the insulation or adjacent to a conductor. Note: Corona is a partial discharge that occurs in gaseous media around conductors that are remote from solid or liquid insulation. This term is not to be used as a general term for all forms of partial discharges.
Discharge (disruptive)	The passage of an arc following the breakdown of the insulation. Note: Discharges are often described as arcing, breakdown or short circuits. More specific terms are sparkover (discharge through oil), puncture (discharge through solid insulation), flashover (discharge at the surface of solid insulation), tracking (progressive degradation of solid insulation by local discharges to form conducting or partially conducting paths) and sparking (local dielectric breakdowns of high ionisation density or small arcs).
Thermal fault	Excessive temperature rise in the insulation. Note: More specific terms are overheating and hotspots.
Typical values of gas concentrations	Gas concentrations normally found in equipments in service which have no symptom of failures, typically refer to values of which 90% or 95% of gas concentrations fall below.

3.4.1 Dörnenburg Ratios

In 1970, Dörnenburg differentiated between faults of thermal or electrical origin by plotting successively the ratio of CH_4/H_2 against C_2H_2/C_2H_4, C_2H_6/C_2H_2 and C_2H_2/CH_4. It was found that all resulting plots can be partitioned into three sections, which correspond to three kinds of fault: local overheating, weak discharges in gas pockets and other discharges. Limits of foregoing ratios for these faults were summarised and tabled by Dörnenburg, as shown in Table 3.2 [15]. Dörnenburg also recommended typical values of gas concentrations below of which power transformers may be regarded as healthy or operating normally, as shown in Table 3.3 [15].

Table 3.2: Fault diagnosis based on Dörnenburg Ratios

Ratio	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_2}{CH_4}$
Fault				
Local overheating	> 1	< 0.7	> 0.4	< 0.3
Weak discharges in gas pockets	< 0.1	*	> 0.4	< 0.3
All other types of discharges	< 1 > 0.1	> 0.7	< 0.4	> 0.3

* Not indicative of this type of fault, generally not applicable. If this gas ratio is applicable and > 1 this indicates that the discharge is increasing.

Table 3.3: Typical values of gas concentrations

Key Dissolved Gas	Concentration (PPM)
H_2	200
CH_4	50
C_2H_6	15
C_2H_4	60
C_2H_2	15
CO	1000
CO_2	11000

The applicability of Dörnenburg Ratios are dictated by the following guidelines [15]:

- A single ratio can be used in the diagnosis only if the concentration of one of the two gases is twice as high as the limiting value quoted in Table 3.3.
- Several ratios can be used together in the diagnosis if at least one of the first and second ratios can be used alone according to the first rule above and at least one of these gases which concentrations are formed by the other ratio, exceeds the limiting value quoted in Table 3.3.
- The additional mixture ratios C_2H_6/C_2H_2 and C_2H_2/CH_4 are acceptable as additional confirmation of the diagnosis if at least one of each pair of gases in the ratio exceeds the concentration quoted in Table 3.3.
- This method of diagnosis may be used only with extreme caution if the gases dissolved in the oil originate from a fault which has no longer been present for some considerable time; various gases of degradation travel at various speeds towards the surface of the oil in the expansion tank and escape into the atmosphere. When the generation of gases stops this distorts the gas ratio.
- CO and CO_2 are found typically as a result of degradation of the solid insulation and are not used in the ratios. In the case of gradual faults, it is generally not possible to decide whether and to what degree the solid insulation has been attacked.
- In the case of transformers with a gas cushion (e.g. nitrogen cushion) above the oil level, limiting values quoted in Table 3.3 for the ratios can be applied only to a limited extent. If volumes of oil and gas cushion are known quantities, the applicable limiting values can be calculated.

3.4.2 Rogers Ratios

This interpretation method was introduced to the Central Electricity Generation Board (CEGB), England and Wales in the 1970's as a method based on Halstead's hypothesis on the generation of hydrocarbon gases in insulation oil [31]. It is based on four gas ratios: CH_4/H_2, C_2H_6/CH_4, C_2H_4/C_2H_6 and C_2H_2/C_2H_4; codes are

assigned to gas ratios according to their limits, as shown in Table 3.4. The diagnosis of a fault is accomplished based on various combinations of codes, as shown in Table 3.5.

Table 3.4: Assignment of codes based on limits of ratios

Ratio	Limit		Code
$\frac{CH_4}{H_2}$	Not greater than 0.1	(≤ 0.1)	5
	Between 0.1 and 1.0	($> 0.1, < 1$)	0
	Between 1.0 and 3.0	($\geq 1, < 3$)	1
	Not less than 3.0	(≥ 3)	2
$\frac{C_2H_6}{CH_4}$	Less than 1.0	(< 1)	0
	Not less than 1.0	(≥ 1)	1
$\frac{C_2H_4}{C_2H_6}$	Less than 1.0	(<1)	0
	Between 1.0 and 3.0	($\geq 1, < 3$)	1
	Not less than 3.0	(≥ 3)	2
$\frac{C_2H_2}{C_2H_4}$	Less than 0.5	(< 0.5)	0
	Between 0.5 and 3.0	($\geq 0.5, < 3$)	1
	Not less than 3.0	(≥ 3)	2

Table 3.5: Fault diagnosis based on combinations of codes

$\frac{CH_4}{H_2}$ **Code 1**	$\frac{C_2H_6}{CH_4}$ **Code 2**	$\frac{C_2H_4}{C_2H_6}$ **Code 3**	$\frac{C_2H_2}{C_2H_4}$ **Code 4**	**Diagnosis**
0	0	0	0	Normal deterioration (no fault)
5	0	0	0	Partial discharges
1/2 *	0	0	0	Slight overheating (below 150 °C)
1/2 *	1	0	0	Overheating (150 – 200 °C)
0	1	0	0	Overheating (200 – 300 °C)
0	0	1	0	General conductor overheating
1	0	1	0	Winding circulating currents
1	0	2	0	Core and tank circulating currents, overheated joints
0	0	0	1	Flashover without power follow through
0	0	1/2 *	1/2 *	Arc with power follow through
0	0	2	2	Continuous sparking to floating potential
5	0	0	1/2 *	Partial discharges with tracking

* Gas ratio of 1 or 2.

3.4.3 IEC Ratios

This interpretation approach is introduced by the IEC 60599 Standard [19,20], which was published in 1978 [19] and was revised in 1999 [20]. For simplicity, the former approach is referred to as the IEC-1978 Ratios and the latter as the IEC-1999 Ratios. Generally, the interpretation approach as employed by the IEC-1978 Ratios is similar to that of Rogers Ratios. However, only three gas ratios are considered, i.e. C_2H_2/C_2H_4, CH_4/H_2 and C_2H_4/C_2H_6; C_2H_6/CH_4 is omitted since it only indicates limited temperature range of degradation and does not assist in further identifying the fault [19]. This method of DGA interpretation is illustrated in Table 3.6. The detection of cellulose degradation due to the fault, which is based on CO_2 and CO, is also recommended by the standard, as outlined in Table 3.7.

Table 3.6: Fault diagnosis based on IEC-1978 Ratios

	Ratio of Characteristic Gases	**Code of Ratio**		
		C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
	< 0.1 0.1 – 1 1 – 3 > 3	0 1 1 2	1 0 2 2	0 0 1 2
Case	**Diagnosis**			
1	No fault	0	0	0
2	Partial discharges of low energy density	Not significant	1	0
3	Partial discharges of high energy density	1	1	0
4	Discharges of low energy	1 → 2 *	0	1 → 2 *
5	Discharges of high energy	1	0	2
6	Thermal fault of temperature (< 150 °C)	0	0	1
7	Thermal fault of low temperature range (150 °C to 300 °C)	0	2	0
8	Thermal fault of medium temperature range (300 °C to 700 °C)	0	2	1
9	Thermal fault of high temperature range (> 700 °C)	0	2	2

* Gas ratio increases from 1 to 2.

Table 3.7: Detection of cellulose degradation in IEC-1978 Ratios

Detection of Cellulose Degradation due to Fault
1. Any case in which CO_2/CO is below about 3 or above about 11 should be regarded as perhaps indicating a fault involving cellulose, if other key gases also indicating excessive oil deterioration.
2. When there is suspicion that a fault might have involved cellulose, the significance of these gases must be considered against the possibility of both gases may have present in moderately high concentration before fault occurred.

The revision of IEC 60599 Standard was accomplished in 1999, resulting in a new approach for interpretation of dissolved gases [20]. The approach of interpretation has been changed from code-assignment to dealing directly with limits of ratios, as shown in Table 3.8. In addition, a method for the detection of cellulose degradation due to a fault is also included, as outlined in Table 3.9. Lastly, recommended procedures for implementation of IEC-1999 Ratios are also outlined in the standard, as illustrated in Figure 3.5.

Table 3.8: Fault diagnosis based on IEC-1999 Ratios

Case	Characteristic Fault	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
PD	Partial discharges	NS [a]	< 0.1	< 0.2
D1	Discharges of low energy	> 1	0.1 – 0.5	> 1
D2	Discharges of high energy	0.6 – 2.5	0.1 – 1	> 2
T1	Thermal fault (T < 300 °C)	NS [a]	> 1 but NS [a]	< 1
T2	Thermal fault (300 °C < T < 700 °C)	< 0.1	> 1	1 – 4
T3	Thermal fault (T > 700 °C)	< 0.2 [b]	> 1	> 4

a. Not significant whatever the value.
b. An increasing value of C_2H_2 may indicate that the hotspot temperature is higher than 1000 °C.
c. The above ratios are significant and should be calculated only if at least one of the gases is at a concentration and at a rate of gas increase above typical values.

Table 3.9: Detection of cellulose degradation in IEC-1999 Ratios

Detection of Cellulose Degradation due to Fault
Incremental (corrected) CO_2/CO ratios less than 3 are generally considered as an indication of probable paper involvement in a fault, with some degree of carbonisation

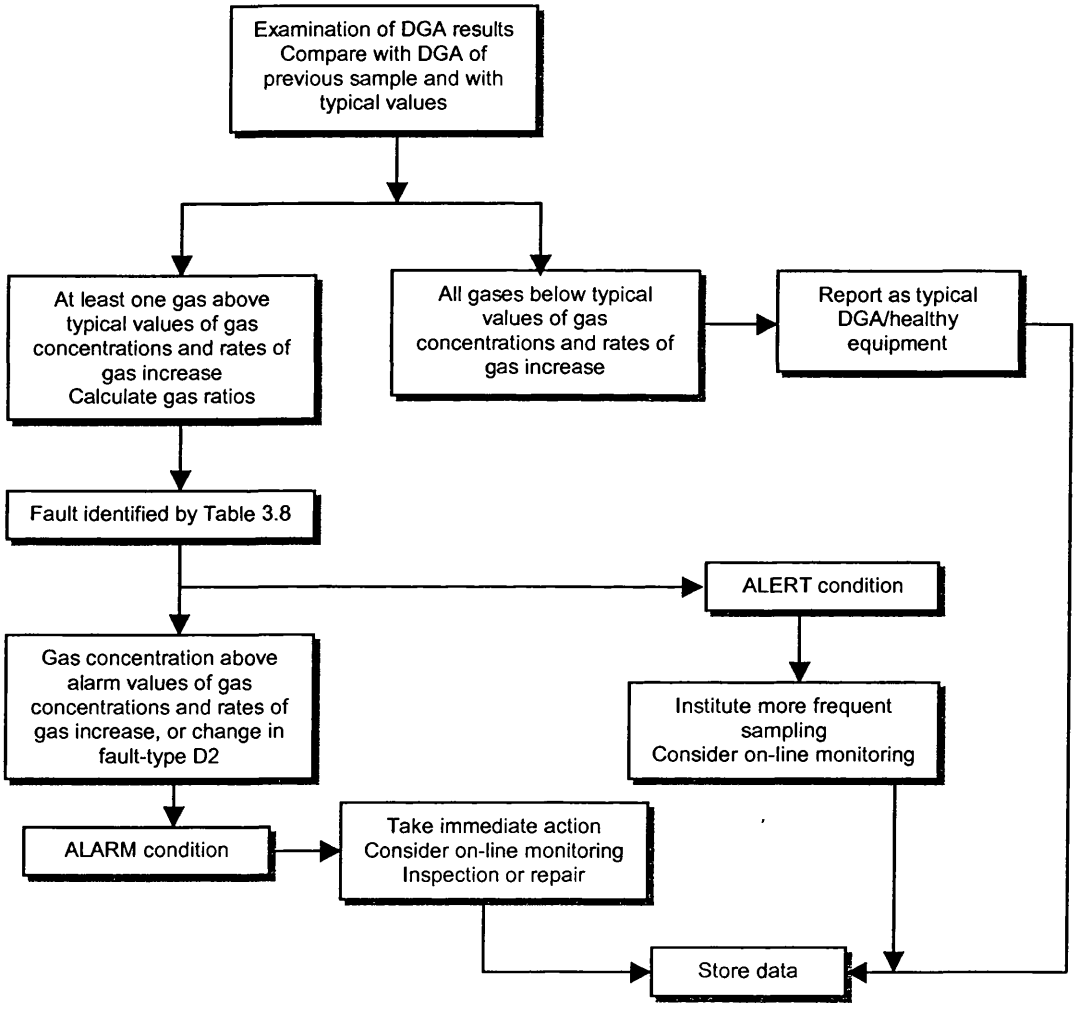
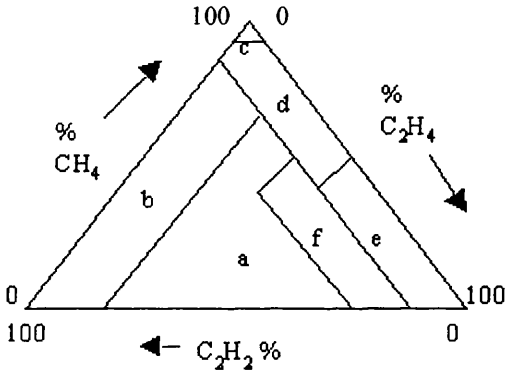


Figure 3.5: Recommended procedures for implementation of IEC-1999 Ratios

3.4.4 Duval Triangle

Duval's method of DGA interpretation was founded on his experience on Hydro-Québec's power transformers [17], in which a triangular diagram has been invented for fault diagnosis based on the concentration of CH_4, C_2H_4 and C_2H_2. Duval's triangular diagram comprises of six fault regions, which include corona discharges, low and high energy arcing, and hotspots of various temperatures, as illustrated in Figure 3.6. These fault regions were identified by plotting individual points onto the diagram [17]; each of these points represents the percentage-concentration of CH_4, C_2H_4 and C_2H_2. The type of fault corresponding to each point has been identified by the kind of damage observed afterwards during the inspection of the faulted equipment.



a: High-energy arcing
b: Low-energy arcing, tracking
c: Corona discharges
d: Hotspots (T < 200 °C)
e: Hotspots (200 °C < T < 400 °C)
f: Hotspots (T > 400 °C)

With x = [C_2H_2]; y = [C_2H_4]; z = [CH_4]

$$\%C_2H_2 = \frac{100x}{x+y+z}$$

$$\%C_2H_4 = \frac{100y}{x+y+z}$$

$$\%CH_4 = \frac{100z}{x+y+z}$$

Figure 3.6: Duval Triangle (1980 edition)

Duval revised the triangular diagram in 1993 [18], with improvements such as the ability to detect simultaneous occurrence of discharges and thermal fault (TF), as illustrated in Figure 3.7. Besides, the location of TF regions has also been amended.

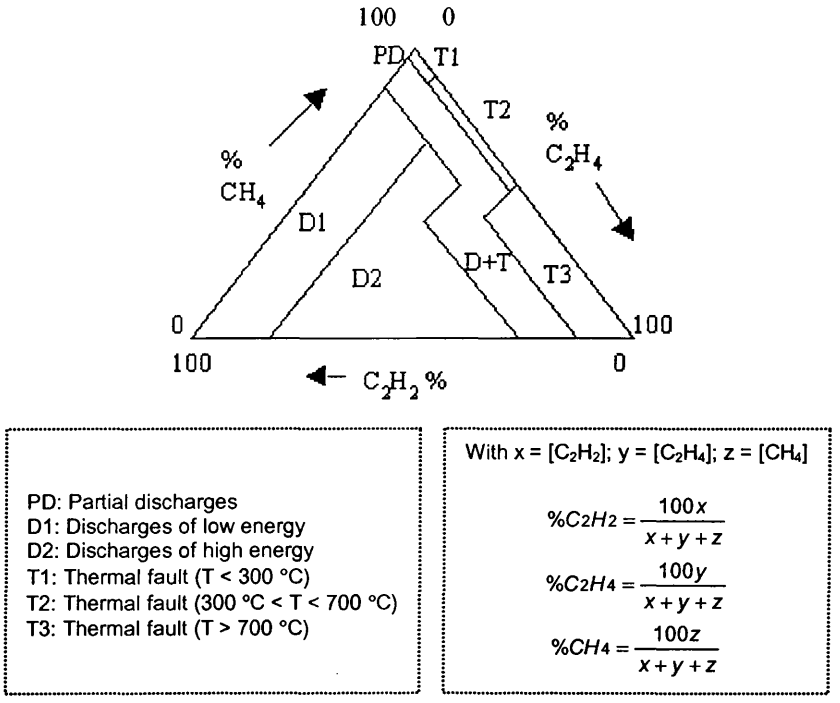


Figure 3.7: Duval Triangle (1993 edition)

3.4.5 CIGRÉ Methods

CIGRÉ methods for the interpretation of DGA data were recommended by the CIGRÉ Task Force 15.01.01 in 1999 in order to fulfil the following aims [21]:

- Proposal of typical values for normal concentrations of key gases as mentioned in the IEC 60599 Standard (1999 edition).
- Reduction of the number of faults listed in the interpretation table of IEC 60599 Standard (1999 edition).
- Precise information at which level of gases, the method of interpretation is applicable.
- Definition of realistic values for CO_2/CO ratio to improve the diagnoses on cellulose insulation.

- Choice of different ratios for the discrimination of faults and determination of the limits for different types of equipment.
- Application of the method to equipment other than power transformers.

Consequently, two DGA interpretation approaches were recommended, which are based on key-gas ratios and key-gas concentrations, respectively. The former approach is based on five key-gas ratios: C_2H_2/C_2H_6, H_2/CH_4, C_2H_4/C_2H_6, CO_2/CO and C_2H_2/H_2, as illustrated in Table 3.10. On the other hand, the latter approach is based on four categories of dissolved gases: C_2H_2, H_2, sum of hydrocarbon gases and sum of carbon oxides, as illustrated in Table 3.11.

Table 3.10: Fault diagnosis based on CIGRÉ's key-gas ratios

Key-gas Ratio	Associated Fault and Limit of Ratio
$\frac{C_2H_2}{C_2H_6}$	Discharges if ratio > 1
$\frac{H_2}{CH_4}$	Partial discharges if ratio > 10
$\frac{C_2H_4}{C_2H_6}$	Thermal fault if ratio > 1
$\frac{CO_2}{CO}$	Cellulose degradation due to overheating if ratio > 10 Cellulose degradation due to electrical fault if ratio < 3
$\frac{C_2H_2}{H_2}$	Discharges in in-tank tap-changer if $C_2H_2 \geq 30$ and ratio ≥ 2

Table 3.11: Fault diagnosis based on CIGRÉ's key-gas concentrations

Key Gases	Key-gas Concentration (PPM)	Suspect of Indication
C_2H_2	> 20	Power discharges
H_2	> 100	Partial discharges
ΣC_xH_y	> 1000 > 500	Thermal fault if up to $\Sigma C_1, C_2, C_3$ – Hydrocarbons if up to $\Sigma C_1, C_2$ – Hydrocarbons
ΣCO_x x = 1, 2	> 10000	Cellulose degradation

3.5 AI-Based Fault Diagnosis Approaches

Attempts have been made to utilise AI techniques for fault diagnosis of power transformers based on the DGA data of power transformers. The main aim of these approaches is to resolve some identified weaknesses of conventional DGA interpretation schemes and to improve the accuracy of diagnosis. Earlier AI techniques used are, for example, expert systems (ESs) [43-47], which are capable of replicating the expert judgement for fault diagnosis. On the other hand, fuzzy logic (FL) [43-46] was also utilised in order to resolve the ambiguity in the diagnosis process.

Another breakthrough is the application of supervised neural-networks (NNs) [32-37], which have the ability of learning the inherent correlation between the composition of dissolved gases and faults in transformers. In addition, unsupervised-NN was also suggested for analysis of dissolved-gas information so as to unearth useful knowledge from the data [38-41]. On the other hand, application of some new algorithms such as evolutionary programming (EP) [45] and self-organising polynomial network (SOPN) [42] were also reported.

Generally, the foregoing AI techniques can be categorised into single-AI approaches and hybrid-AI approaches. The former approaches involve the utilisation of one AI techniques while the latter approaches involve the combinatory application of two or more AI techniques.

3.5.1 Single-AI Approaches

The most common AI technique within this category is the supervised-NN. A set of input-output samples is required for the supervised training process. The inputs are, for example, key gases (e.g. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2) and the outputs are, for example, fault conditions (e.g. discharges, PD and TF) that have been identified

either through the application of conventional DGA interpretation schemes or through actual inspection on faulted transformers.

In 1993, Bhattacharyya et. al. [32] applied a simple feed-forward NN trained with back-propagation algorithm for detecting thermal and arcing faults. The inputs used are H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2, O_2, N_2, sum of all gases and total combustible gases (TCG). Training samples were taken from post mortem data and were carefully selected so that various operating conditions were well represented. It was reported that more accurate diagnoses can be obtained if compared to Rogers Ratios and Dörnenburg Ratios.

In 1996, Zhang et. al. [33] developed a two-step NN approach, where two NNs were trained for fault diagnosis and detection of cellulose degradation respectively, as illustrated in Figure 3.8. It was found that CO_2 and CO are not needed as inputs for fault diagnosis, and a single output that indicates whether cellulose was involved in a fault is sufficient for the detection of cellulose degradation. In addition, Zhang et. al. also indicated in latter publication [34] that higher diagnosis accuracy could be achieved if gas generation rates were included as inputs to the NN.

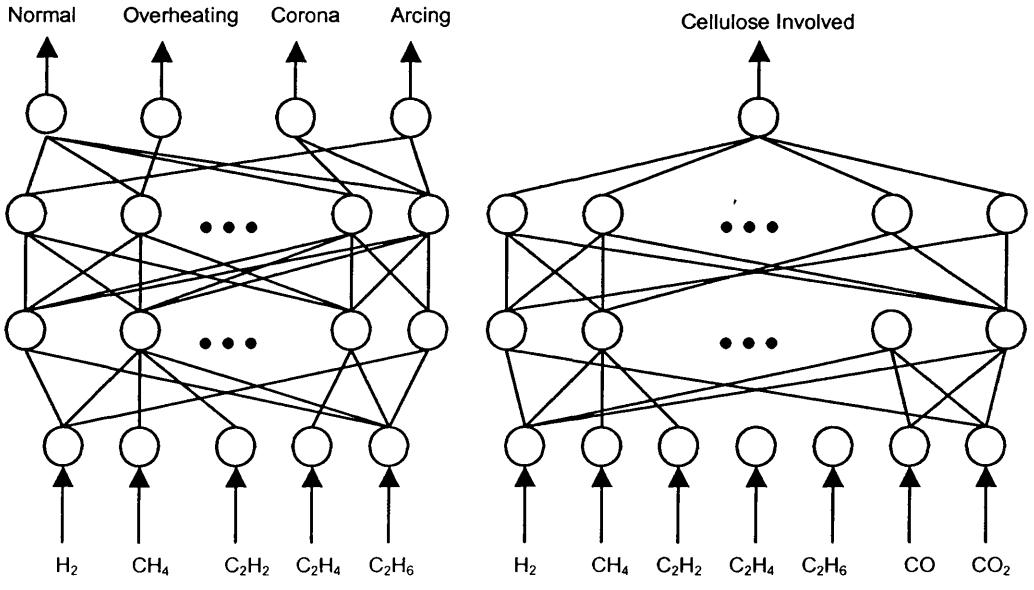


Figure 3.8: Two-step supervised-NN approach

In 1997, Tu et. al. [35] utilised an approach of subdivided NNs. Three NNs that correspond to key gas concentrations, IEC-1978 Ratios and Rogers Ratios were trained independently; fault diagnoses as given by these NNs were combined through a process known as "synthetic analysis" to arrive at a final decision on fault diagnosis. However, Tu et. al. pointed out that the selection of suitable training cases and their quantities would affect the diagnosis capability of supervised-NNs.

Supervised-NN was also used by Venegas et. al. [36] in 1997 for fault diagnosis based on the Japanese ECRA method. Thirteen characteristic patterns of gaseous composition were used as inputs to the NN, which was trained to detect different types of fault as specified in the method. It was reported that 92% of diagnosis accuracy could be achieved through the use of this approach. Finally, a multi-resolution approach was developed by Gao et. al. [37] in 1998, where different supervised-NNs were connected to form a decision tree. Fault diagnosis was accomplished from top to bottom, where a more detailed fault type was detected in the lower layer of the decision tree.

Besides, an unsupervised-NN known as the self-organising map (SOM) was applied by Esp et. al. [38-41] for explorative analysis on selected set of DGA records. It was reported that some interesting patterns were revealed that can be associated with certain incipient faults of power transformers. This unsupervised-NN approach will be further explored in a systematic and detailed manner in the next few chapters of this thesis. In fact, the research project as reported in this thesis is a result of collaboration with Esp et. al. in order to further investigate the foregoing approach.

In 1998, Yang et. al. [42] applied a novel data-driven approach, which is known as the self-organising polynomial network (SOPN), for fault diagnosis of power transformers, as shown in Figure 3.9. The SOPN is similar to supervised-NN topologically but the number of layers, number of nodes, connection weights and the transfer function for each node must be determined through iterative training. Therefore, the non-linear nature of the data is learned through a cascaded architecture of simple low-order polynomial functions. In addition, it was reported that the

training of SOPN is seven times faster than supervised NNs and more accurate diagnoses were also obtained.

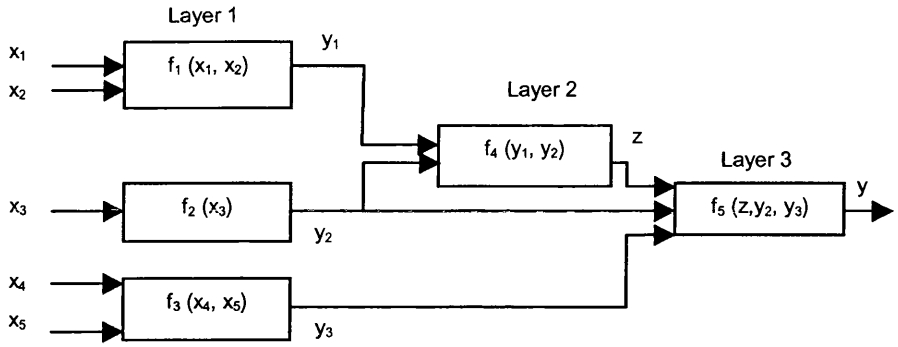


Figure 3.9: Self-organising polynomial network

3.5.2 Hybrid-AI Approaches

Other AI techniques applied for fault diagnosis of power transformers are of hybrid nature, such as the fuzzy expert system (FES) [43, 44], fuzzy evolutionary programming (FEP) [45], fuzzy neural network (FNN) [46] and combined application of ES and NN [47]. The advantage of employing hybrid approaches is that the strength of each individual technique can be effectively combined and hence more accurate fault diagnoses can be achieved.

A FES, as illustrated in Figure 3.10, was developed by Lin et. al. in 1993 [43]. The FES was implemented using if-then rules and fuzzy logic (FL) was introduced to resolve the inherent uncertainties in normality thresholds, gas ratios and key gas concentrations. The knowledge-base of the FES incorporates not only popular DGA interpretation schemes such as the Dörnenburg Ratios, key gas concentrations and IEC-1978 Ratios, but also synthetic expertise and heuristic maintenance rules based on expert experiences.

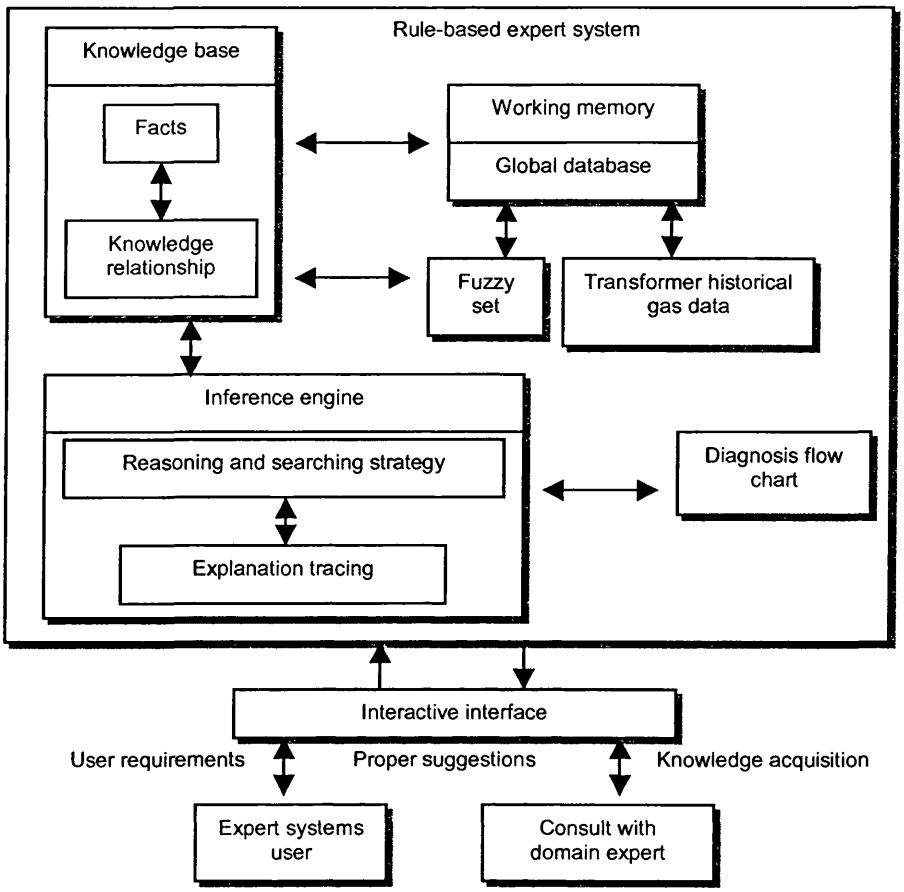


Figure 3.10: Fuzzy expert system

However, only a fairly simple form of fuzzy concepts was implemented by Lin et. al. and the more general framework associated with fuzzy measures and bodies of evidence were not pursued [44]. Consequently, a more general approach, known as the fuzzy information theory, was proposed by Tomsovic et. al. [44] in 1993 in order to systematically manage uncertainties that arise from different DGA interpretation schemes. In the FES developed by Tomsovic et. al., each DGA interpretation scheme was represented by several fuzzy rules; conflicts that arise between rules were resolved using fuzzy information theory to find the most consistent solution. Lastly, diagnoses as given by various schemes were combined to arrive at a final decision, in which higher weights were attached to more certain diagnoses.

Although the introduction of fuzzy concepts greatly improves the diagnosis accuracy of expert system, membership functions of fuzzy subsets are either determined empirically or basically in a trial-and-error manner, where the conventional DGA interpretation schemes are to be implicitly followed. Huang et. al. [45] proposed a novel approach known as the fuzzy evolutionary programming (FEP) in 1997, whereby conventional DGA interpretation schemes were used to construct the preliminary framework of the fuzzy system, and an evolutionary programming (EP) based optimisation algorithm was employed to further modify the fuzzy IF-THEN rules and simultaneously adjusting the membership function of the fuzzy subsets. The overall FEP approach is illustrated in Figure 3.11. Consequently, the cumbersome process of manually adjusting the fuzzy rules and membership functions, as experienced by Lin et. al. [43] and Tomsovic et. al. [44] can be avoided altogether.

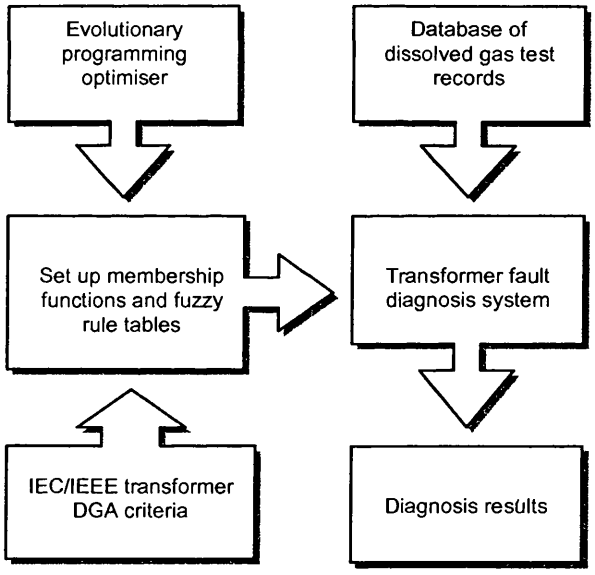


Figure 3.11: Fuzzy evolutionary programming

In 1997, Tomsovic et. al. [46] attempted to compare the diagnosis accuracy of three AI approaches: fuzzy information approach as reported in [44], supervised-NN, and a hybrid approach known as the fuzzy neural network (FNN). For the FNN approach, developed fuzzy relations in the fuzzy information approach were used for

generating inputs to the supervised-NN. However it was reported that the FNN produces similar diagnosis accuracy to that of supervised-NN; it also failed to achieve more accurate results when compared with the fuzzy information approach [44].

A combined NN and ES approach was developed and reported by Wang et. al. [47] in 1998, as illustrated in Figure 3.12. The knowledge-base of the expert system integrates IEEE Standard [48], IEC Standard [19] and human expertise to ensure that the additional knowledge was not abandoned when insufficient data was available for the supervised-NN training. On the other hand, the supervised-NN could acquire new experiences through incremental training from newly obtained data samples. An optimisation mechanism was applied to combine respective outputs from ES and supervised-NN to arrive at the final diagnosis, accompanied by maintenance recommendations. This approach was reported to have better performance when compared with supervised-NN and ES utilised individually.

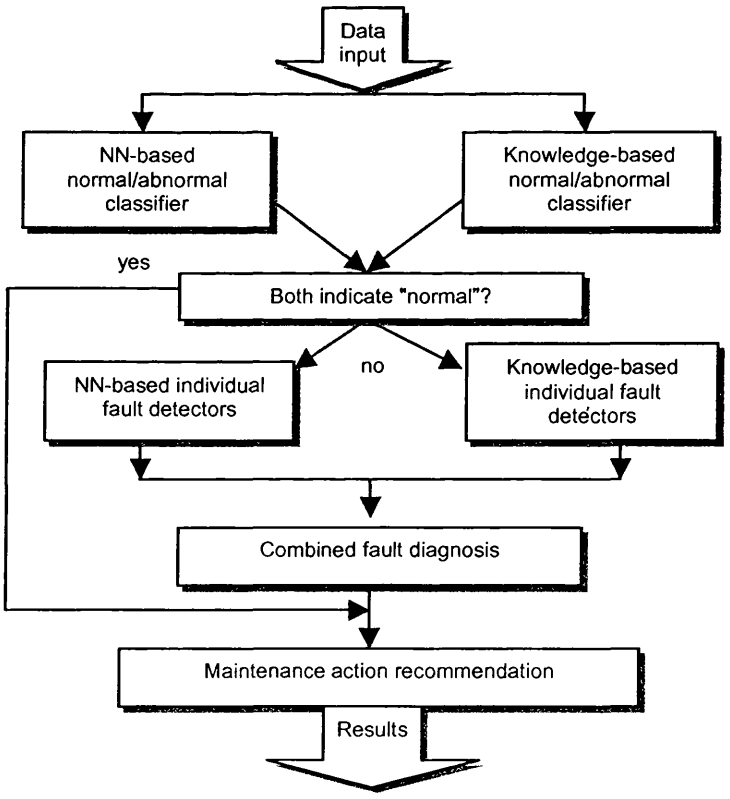


Figure 3.12: Combined expert system and supervised NN approach

3.6 Effectiveness of Current Fault Diagnosis Approaches

As briefly explained in Sections 3.4 and 3.5, current fault diagnosis approaches can be classified into conventional DGA interpretation schemes [15-21] and new AI-based approaches [32-47]. The former approaches are based on the key-gas ratios and key-gas concentrations while the latter approaches are dependent on the application of AI techniques for enhancing the accuracy and consistency of the fault diagnosis.

The conventional DGA interpretation schemes have been known to contain several weaknesses. Firstly, ambiguity still exists on the types of key-gas ratios to be considered and on the credibility of the suggested ratio limits; power utilities may have to select or adapt any of the well-established schemes heuristically. Even so, there is no guarantee that the chosen or adapted DGA interpretation scheme is best suited for accomplishing the task of fault diagnosis.

Secondly, the empirical nature of conventional DGA interpretation schemes has led to discrepancies in interpretation; application of various interpretation schemes on a set of DGA records may result in diverse interpretation of transformer condition, thereby causing confusion among power utilities.

Thirdly, interpretation of transformer condition is sometimes impossible to achieve owing to the inability of some interpretation schemes to cover all possible combinations of gas ratios or codes. Consequently, the interpretation of a DGA record may have to depend on human expert, who may instigate even more confusion since each expert may have his/her own opinion on what is happening inside the transformer.

Finally, no systematic attempt has been made in these schemes to actually "learn" from the hidden information contained within the historical DGA database, which

may contain a lot of useful information regarding series of events or incidences that have occurred inside the power transformers.

On the other hand, several AI-based approaches have been applied for resolving some identified weaknesses of conventional DGA interpretation schemes, such as the ambiguity of ratio limits and discrepancies in DGA interpretation. Subsequently, improvements in fault diagnosis were reported. Despite the advantages of utilising these AI-based approaches, they are not entirely ideal and do suffered from several weaknesses.

To begin with, the application of supervised-NNs requires either actual fault cases or conventional DGA interpretation schemes for their modelling. However, the acquisition of actual fault cases is difficult since it is too costly to disconnect and dismantle a particular power transformer for the purpose of investigating a suspected fault since the transformer concerned may seem to be operating normally despite the increase in certain key gases. Consequently, there might not have a lot of "good" fault cases for the training of supervised-NNs. Therefore, the trained NNs might not be able to learn well and subsequently incapable of generalising towards some new cases that are presented to them.

Furthermore, as for those supervised-NNs that are utilising conventional DGA schemes for generating targets for the training inputs, they will inevitably inherit some of the inherent weaknesses of these schemes such as the ambiguity of suggested ratio limits and gas-ratios. Therefore, the trained NNs in this case would only provide limited generalisation capability owing to their dependency on conventional schemes.

Besides, hybrid-AI approaches such as FNN, FES and combined ES and NN are more effective due to the fact that FL or NN is used to tackle the ambiguity of conventional DGA interpretation schemes, which are integrated into foregoing approaches, and expert experiences are incorporated to improve the credibility of diagnosis. However, owing to the incorporation of conventional schemes and expert

experiences, too many uncertainties are introduced to these approaches and would lead to the lack of confidence on final diagnoses if these uncertainties were not managed appropriately.

It is apparent from the foregoing that there is a need for research into new approaches, which ideally do not depend on actual fault cases or conventional DGA interpretation schemes for development and modelling processes while offering improved accuracy and confidence in fault diagnosis. With this background, a novel approach is introduced in this thesis that is based on the application of an unsupervised-NN known as the SOM. The feasibility and ability of this proposed approach for addressing the afore-mentioned issues and achieving foregoing aims will be investigated in subsequent chapters of this thesis.

3.7 Summary

Various aspects of DGA have been presented in this chapter, which include the related theory, procedures of the DGA test and a literature survey on conventional DGA interpretation schemes and new AI-based fault diagnosis approaches. It has been pointed out that the DGA approach is based on the mechanisms of formation of gases in insulation oil during the onset of faults. Besides, various procedures of the DGA test have been briefly illustrated. In addition, a discussion on the effectiveness of currently available fault diagnosis approaches have been presented and it is concluded that a new approach is clearly needed for improved fault diagnosis in a more systematic and confident manner.

CHAPTER 4

IMPLEMENTATION OF CONVENTIONAL DGA INTERPRETATION SCHEMES

4.1 Introduction

Conventional DGA interpretation schemes as described in Section 3.4 were implemented using the Matlab programming language. Modifications and improvements were made so as to maintain the overall uniformity and to mimic, as close as possible, actual circumstances to which these schemes are applied. In addition, a comparative study on the implemented schemes is also presented in this chapter, with the aim of investigating the effectiveness of these schemes for the interpretation of the DGA data of power transformers.

4.2 Important Aspects of the Implementation

Various aspects have to be considered before these DGA interpretation schemes are implemented. Firstly, a distinction must be established between the diagnosis of incipient faults and detection of cellulose degradation due to these faults. In fact, DGA is very efficient for detecting incipient faults but not very useful when it comes to determining the degree of cellulose degradation due to these faults. Therefore, main emphasis is placed herein on the ability of the implemented schemes to detect electrical faults (EF), i.e. discharges and partial discharges (PD), and thermal faults (TF), e.g. hotspots and overheating.

Secondly, the terminology of incipient faults has to be standardised so as to avoid confusion in the later stages. EF is thus used as a general term for discharges and PD. Specifically, the term "discharges" is used for referring to sparkover, puncture, flashover, tracking and sparking; PD is used as a general term for partial dielectric breakdown that includes corona, which is a specific form of PD in a gaseous medium, and finally, TF is used for referring to hotspots and other overheating problems. Definitions for the foregoing terms have been presented in Table 3.1.

Thirdly, typical concentrations of several key dissolved gases have to be determined, which represent concentrations of these gases which are usually found in insulation oil of power transformers that are operating normally. Since typical concentrations are explicitly stated in Dörnenburg Ratios and CIGRÉ key-gas concentrations method, they will be used as typical values for these approaches. As for the other DGA interpretation schemes, typical concentrations have to be calculated from a historical DGA database, which contains a total of 14943 DGA records from around 600 power transformers of the National Grid Company (NGC), UK. The determination of typical concentrations is based on the assumption that 90% of recorded dissolved-gas concentrations correspond to power transformers that are operating normally and the remaining 10% indicate the possibility of faults or abnormal circumstances within the power transformers. Typical concentrations based on the 90% limit can be determined via the plotting of cumulative histograms. Essentially, cumulative histogram for each key dissolved-gas is plotted and a line is then drawn from the 90% cumulative-percentage to arrive at the typical concentration for each gas. The typical concentration for each key dissolved gas was thus calculated via the foregoing approach and is shown in Table 4.1.

Table 4.1: Typical dissolved-gas concentrations based on the 90% limit

Key Dissolved Gas	90% Typical Concentration (PPM)
CO_2	4260
CO	520
H_2	85
CH_4	47
C_2H_6	31
C_2H_4	53
C_2H_2	13

4.3 Implementation of Dörnenburg Ratios

Generally, the applicability of Dörnenburg Ratios is dictated by several constraints, as summarised in Table 4.2. Besides, Table 4.3 illustrates the fault-diagnosis table for Dörnenburg Ratios. In addition, the flow-chart for the implementation of Dörnenburg Ratios is illustrated in Figure 4.1. Note that typical dissolved-gas concentrations as suggested by Dörnenburg (see Table 3.3) were used for the implementation of this method.

Table 4.2: Constraints of Dörnenburg Ratios

Dörnenburg Ratios is Applicable if At Least One of the Following Conditions is Satisfied:
Condition 1: (CH_4 > 100) AND (C_2H_2 > 15 OR C_2H_4 > 60) AND (C_2H_6 > 15 OR C_2H_2 > 15) AND (C_2H_2 > 15 OR CH_4 > 50)
Condition 2: (H_2 > 400) AND (C_2H_2 > 15 OR C_2H_4 > 60) AND (C_2H_6 > 15 OR C_2H_2 > 15) AND (C_2H_2 > 15 OR CH_4 > 50)
Condition 3: (C_2H_2 > 30) AND (CH_4 > 50 OR H_2 > 200) AND (C_2H_6 > 15 OR C_2H_2 > 15) AND (C_2H_2 > 15 OR CH_4 > 50)
Condition 4: (C_2H_4 > 120) AND (CH_4 > 50 OR H_2 > 200) AND (C_2H_6 > 15 OR C_2H_2 > 15) AND (C_2H_2 > 15 OR CH_4 > 50)

Table 4.3: Fault-diagnosis table for Dörnenburg Ratios

Ratio / **Diagnosis**	$\frac{CH_4}{H_2}$	$\frac{C_2H_2}{C_2H_4}$	$\frac{C_2H_6}{C_2H_2}$	$\frac{C_2H_2}{CH_4}$	**Type**
Local overheating	> 1	< 0.7	> 0.4	< 0.3	TF
Weak discharges in gas pockets	< 0.1	Not applicable	> 0.4	< 0.3	PD
All other types of discharges	< 1 > 0.1	> 0.7	< 0.4	> 0.3	Discharges
Other combinations of ratios not covered by Dörnenburg Ratios					No interpretation
One or more ratios are NaNs * (i.e. no record in either one or both gases of a ratio)					Undefined

* NaN: Not-a-Number.

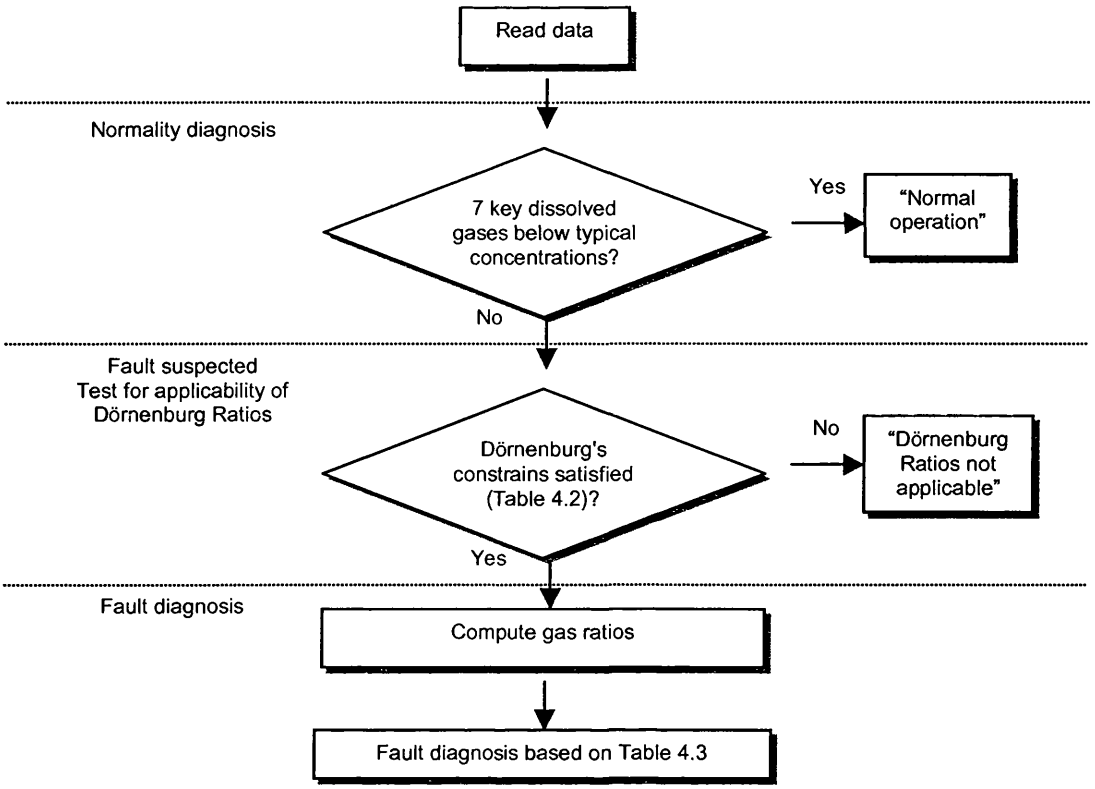


Figure 4.1: Flow-chart for the implementation of Dörnenburg Ratios

4.4 Implementation of Rogers Ratios

The Rogers' method was developed in such a way that at some instances, several combinations of codes could yield identical diagnoses. For example, two combinations of codes are possible for the diagnoses of "slight overheating at below 150 °C", "slight overheating between 150 °C to 200 °C" and "partial discharges with tracking". The implementation of the code-assignment process is illustrated in Table 4.4 and the fault-diagnosis table is shown in Table 4.5. In addition, although the requirement for "normality" test is not explicitly stated in Rogers Ratios, it was implemented so as to mimic actual circumstances in which the method is used. Therefore, the code-combination of [0 0 0 0] in Rogers Ratios was eliminated and typical concentrations of key dissolved gases as listed in Table 4.1 were adopted to cater for such purpose. Finally, the flow-chart for the implementation of Rogers Ratios is illustrated in Figure 4.2.

Table 4.4: Assignment of codes for Rogers Ratios

Code	Ratio	Limit		Assignment of Code
1	$\frac{CH_4}{H_2}$	Not greater than 0.1	(≤ 0.1)	5
		Between 0.1 and 1.0	($> 0.1, < 1$)	0
		Between 1.0 and 3.0	($\geq 1, < 3$)	1
		Not less than 3.0	(≥ 3)	2
		No record in either one or both gases		NaN *
2	$\frac{C_2H_6}{CH_4}$	Less than 1.0	(< 1)	0
		Not less than 1.0	(≥ 1)	1
		No record in either one or both gases		NaN *
3	$\frac{C_2H_4}{C_2H_6}$	Less than 1.0	(< 1)	0
		Between 1.0 and 3.0	($\geq 1, < 3$)	1
		Not less than 3.0	(≥ 3)	2
		No record in either one or both gases		NaN *
4	$\frac{C_2H_2}{C_2H_4}$	Less than 0.5	(< 0.5)	0
		Between 0.5 and 3.0	($\geq 0.5, < 3$)	1
		Not less than 3.0	(≥ 3)	2
		No record in either one or both gases		NaN *

* NaN: Not-a-Number.

Table 4.5: Fault-diagnosis table for Rogers Ratios

Case	Code 1	Code 2	Code 3	Code 4	Diagnosis	Type
1	5	0	0	0	Partial discharges	PD
2	1	0	0	0	Slight overheating (< 150 °C) – 1	TF
	2	0	0	0	Slight overheating (< 150 °C) – 2	TF
3	1	1	0	0	Overheating (150 °C – 200 °C) – 1	TF
	2	1	0	0	Overheating (150 °C – 200 °C) – 2	TF
4	0	1	0	0	Overheating (200 °C – 300 °C)	TF
5	0	0	1	0	General conductor overheating	TF
6	1	0	1	0	Winding circulating current	TF
7	1	0	2	0	Core and tank circulating current	TF
8	0	0	0	1	Flash over without power follow through	Discharges
9	0	0	1	1	Arc with power follow through – 1	Discharges
	0	0	1	2	Arc with power follow through – 2	Discharges
	0	0	2	1	Arc with power follow through – 3	Discharges
10	0	0	2	2	Continuous sparking to floating potential	Discharges
11	5	0	0	1	Partial discharge with tracking – 1	PD
	5	0	0	2	Partial discharge with tracking – 2	PD
12	Other combinations of codes not covered by Rogers Ratios					No interpretation
13	One or more codes are NaNs *					Undefined

* NaN: Not-a-Number.

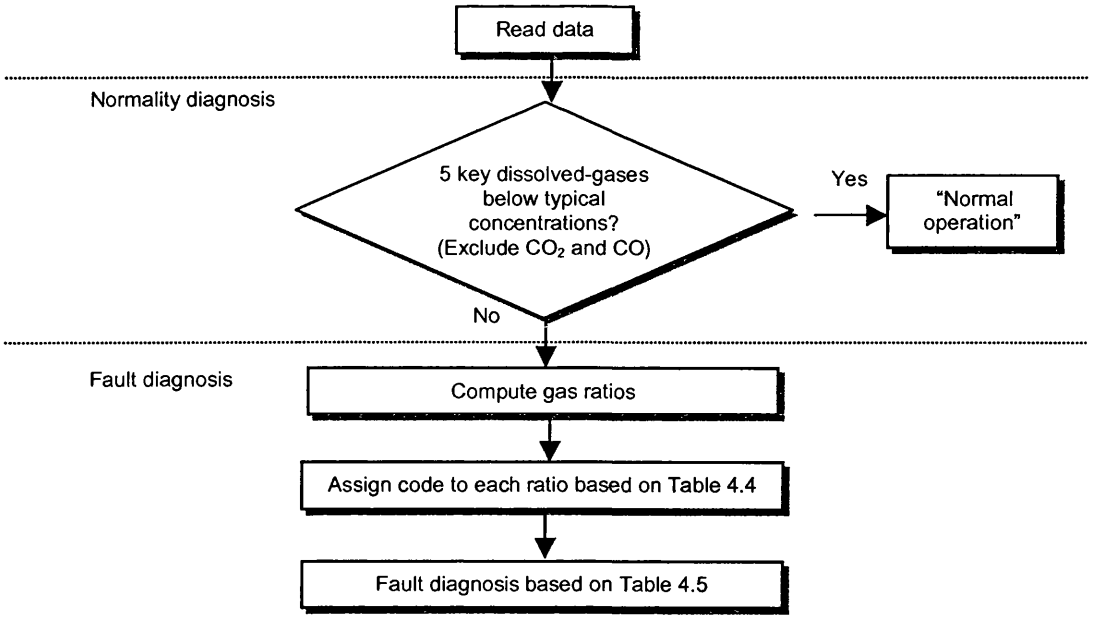


Figure 4.2: Flow-chart for the implementation of Rogers Ratios

4.5 Implementation of IEC-1978 Ratios

In IEC-1978 Ratios, codes are assigned based on respective limits of ratios and fault diagnosis is performed based on various combinations of codes. In addition, there are also instances in which several combinations of codes can lead to identical fault diagnosis. The implementation of the code-assignment process is illustrated in Table 4.6 and the fault-diagnosis table is shown in Table 4.7. Lastly, the flow-chart for the implementation of IEC-1978 Ratios is illustrated in Figure 4.3. Note that typical concentrations as listed in Table 4.1 were used for the "normality" test and the code-combination of [0 0 0] was eliminated to cater for such purpose.

Table 4.6: Assignment of codes for IEC-1978 Ratios

Ratio	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
< 0.1	0	1	0
0.1 – 1	1	0	0
1 – 3	1	2	1
> 3	2	2	2
No record in either one or both gases	NaN *	NaN *	NaN *

* NaN: Not-a-Number.

Table 4.7: Fault-diagnosis table for IEC-1978 Ratios

Case	Diagnosis	$\frac{C_2H_2}{C_2H_4}$ Code 1	$\frac{CH_4}{H_2}$ Code 2	$\frac{C_2H_4}{C_2H_6}$ Code 3
1	Partial discharges of low energy density	0	1	0
2	Partial discharges of high energy density	1	1	0
3	Discharges of low energy – 1 Discharges of low energy – 2 [a] Discharges of low energy – 3 [a]	1 2 2	0 0 0	1 1 2
4	Discharges of high energy	1	0	2
5	Thermal fault of low temperature (< 150 °C)	0	0	1
6	Thermal fault of low temperature (150 °C – 300 °C)	0	2	0
7	Thermal fault of medium temperature (300 °C – 700 °C)	0	2	1
8	Thermal fault of high temperature (> 700 °C)	0	2	2
9	No Interpretation	Other combinations of codes not covered by IEC-1978 Ratios		
10	Undefined	One or more codes are NaNs [b]		

a. Unconvincing combinations of codes, but will be implemented in accordance with the method.
b. NaN: Not-a-Number.

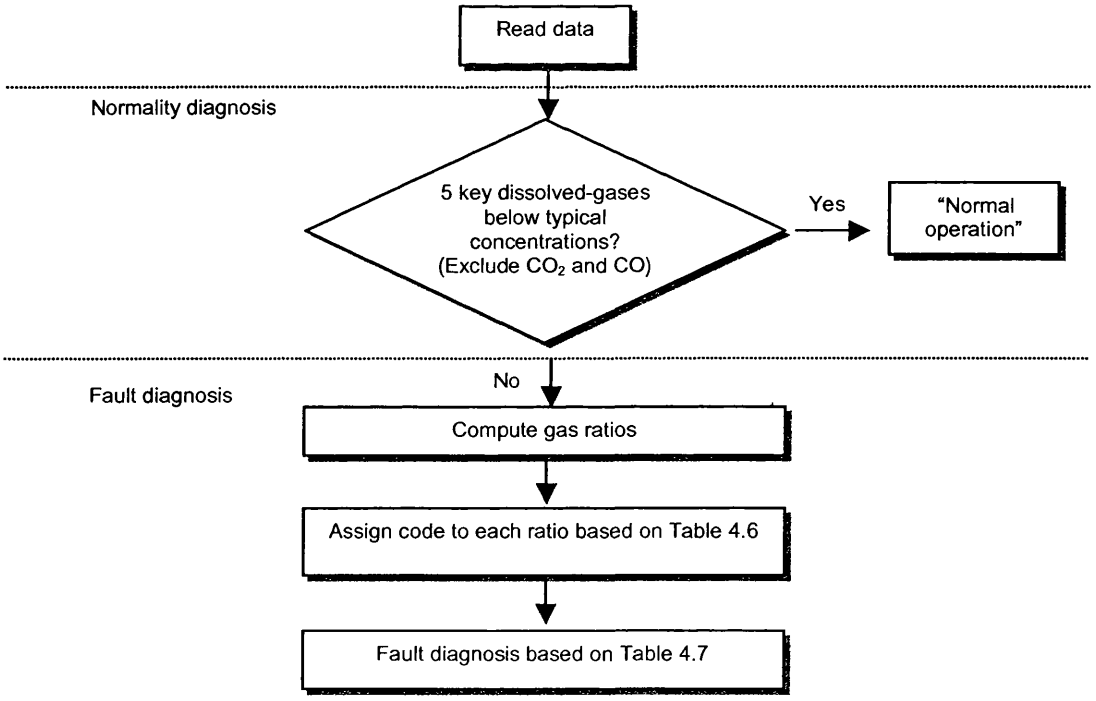


Figure 4.3: Flow-chart for the implementation of IEC-1978 Ratios

4.6 Implementation of IEC-1999 Ratios

The IEC-1999 Ratios is the revised version of the IEC-1978 Ratios. The fault-diagnosis approach has been changed from the assignment of codes to dealing directly with ratios of dissolved gases. In addition, all DGA records are subjected to the "normality" test before being diagnosed for faults. This was done so as to follow the IEC's recommendation on the implementation of this method (see Figure 3.5). Therefore, typical concentrations of dissolved gases that are listed in Table 4.1 were used for that purpose. However, the rates of gas increase are not considered herein owing to the lack of information on the volume of oil in various transformer tanks; this information is needed for the calculation of gas rates. The fault-diagnosis table for IEC-1999 Ratios is illustrated in Table 4.8 and the flow-chart for the implementation of this method is depicted in Figure 4.4.

Table 4.8: Fault-diagnosis table for IEC-1999 Ratios

Case	Diagnosis	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
1	Partial discharges	Not applicable	< 0.1	< 0.2
2	Discharges of low energy	> 1	> 0.1 – ≤ 0.5	> 1
3	Discharges of high energy	> 0.6 – ≤ 2.5	> 0.1 – ≤ 1	> 2
4	Thermal fault (T < 300 °C) – 1	Not applicable	> 1	< 1
	Thermal fault (T < 300 °C) – 2	Not applicable	Not applicable	< 1
5	Thermal fault (300 °C < T < 700 °C)	< 0.1	> 1	> 1 – ≤ 4
6	Thermal fault (T > 700 °C)	< 0.2	> 1	> 4
7	No interpretation	Other combinations of codes not covered by IEC-1999 Ratios		
8	Undefined	One or more ratios are NaNs * (i.e. no record in either one or both gases of a ratio)		

* NaN: Not-a-Number.

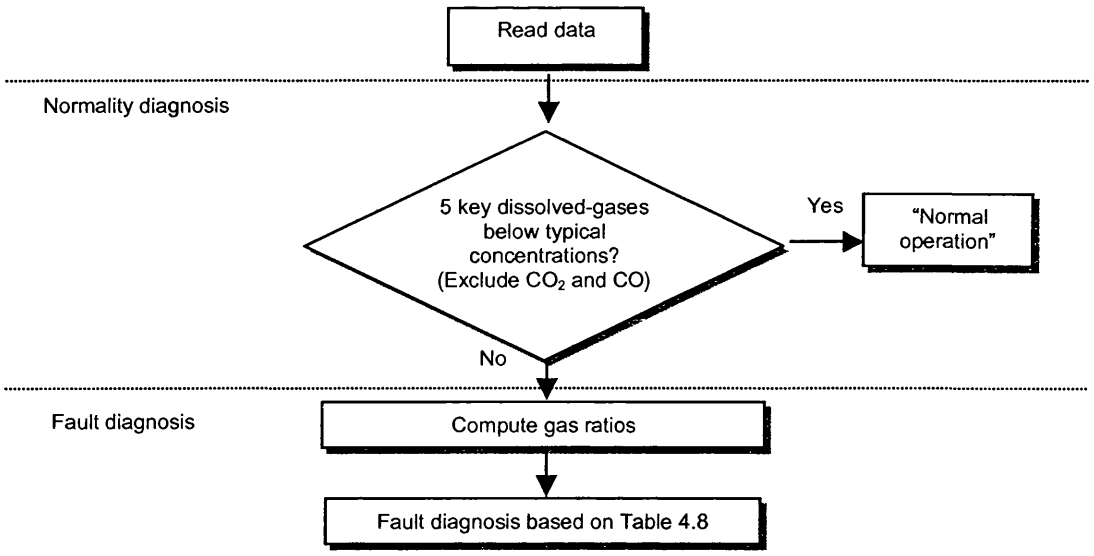


Figure 4.4: Flow-chart for the implementation of IEC-1999 Ratios

4.7 Implementation of Duval Triangle

Although it was not stated that the "normality" test is required before the application of Duval Triangle, this was implemented so as to simulate the actual circumstances in which the method is used; typical concentrations as listed in Table 4.1 were used for that purpose. Two versions of Duval Triangle were implemented, as shown in Tables 4.9 and 4.10 respectively. The flow-chart for the implementation of Duval Triangle is illustrated in Figure 4.5.

Table 4.9: Fault-diagnosis table for Duval Triangle (1980-edition)

Diagnosis	Boundaries for $\%C_2H_2$, $\%C_2H_4$, $\%CH_4$	Type
a: High-energy arcing	• $\%C_2H_2 > 26$ • $26 \geq \%C_2H_2 > 14$ • $\%C_2H_4 > 25$ • $40 \geq \%C_2H_4 > 25$	Discharges
b: Low-energy arcing, tracking	• $\%C_2H_2 > 14$ • $\%C_2H_4 \leq 25$	Discharges
c: Corona discharges	• $\%C_2H_4 > 95$	PD
d: Hot-spots (T < 200 °C)	• $\%C_2H_2 \leq 14$ • $\%C_2H_4 \leq 48$ • $\%CH_4 \leq 95$	TF
e: Hot-spots (200 °C < T <400 °C)	• $\%C_2H_2 \leq 14$ • $\%C_2H_4 > 48$	TF
f: Hot-spots (T > 400 °C)	• $26 \geq \%C_2H_2 > 14$ • $\%C_2H_4 > 40$	TF
One or more %gases is NaN * (i.e. no record in one or more gases of a %gas)		Undefined

* NaN: Not-a-Number.

Table 4.10: Fault-diagnosis table for Duval Triangle (1993-edition)

Diagnosis	Boundaries for $\%C_2H_2$, $\%C_2H_4$, $\%CH_4$
PD: Partial discharges	• $\%CH_4 > 98$
D1: Discharges of low energy	• $\%C_2H_2 > 13$ • $\%C_2H_4 \leq 23$
D2: Discharges of high energy	• $29 \geq \%C_2H_2 > 13$ • $\%C_2H_2 > 29$ • $38 \geq \%C_2H_4 > 23$ • $\%C_2H_4 > 23$
T1: Thermal fault (T < 300 °C)	• $\%C_2H_2 \leq 4$ • $\%C_2H_4 \leq 10$ • $\%CH_4 \leq 98$
T2: Thermal fault (300 °C < T < 700 °C)	• $50 \geq \%C_2H_4 > 10$ • $\%C_2H_2 \leq 4$
T3: Thermal fault (T > 700 °C)	• $\%C_2H_2 \leq 15$ • $\%C_2H_4 > 50$
D+T: Discharges and thermal fault	Except all ranges above
Undefined: One or more %gases is NaN * (i.e. no record in one or more gases of a %gas)	

* NaN: Not-a-Number.

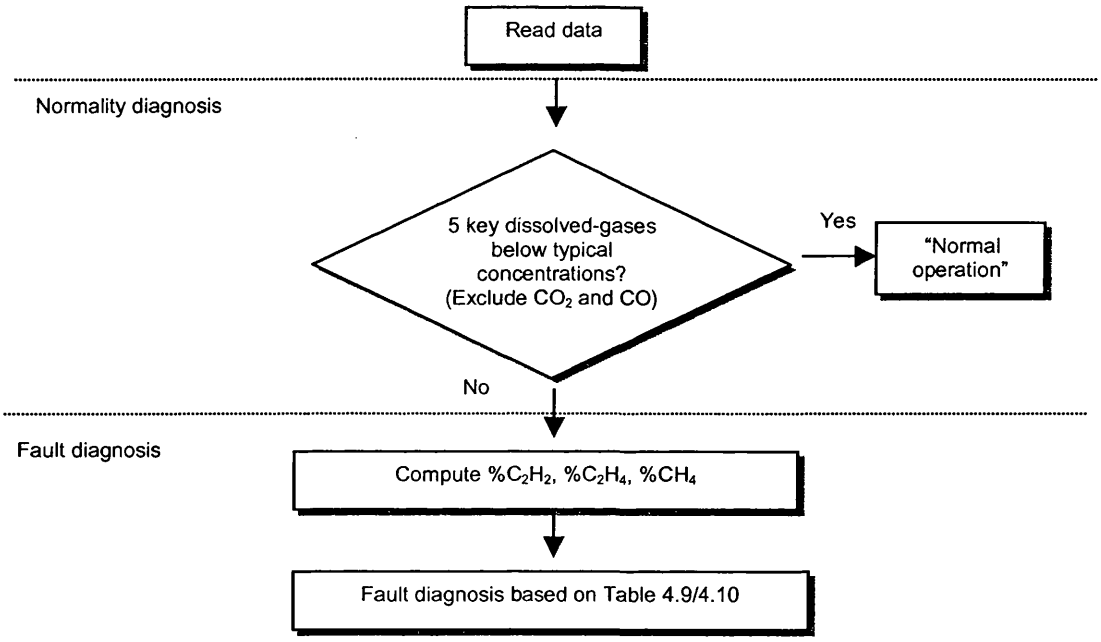


Figure 4.5: Flow-chart for the implementation of Duval Triangle

4.8 Implementation of CIGRÉ Methods

CIGRÉ has recommended the application of key-gas ratios and key-gas concentrations for fault diagnosis of power transformers. Again, the "normality" test was implemented for the key-gas ratios in order to simulate the actual situation in

which the method is used. Fault-diagnosis tables for CIGRÉ's key-gas concentrations and key-gas ratios are shown in Tables 4.11 and 4.12 respectively. In addition, the flow-chart for the implementation of CIGRÉ's key-gas ratios is shown in Figure 4.6.

Table 4.11: Fault-diagnosis table for CIGRÉ's key-gas concentrations

Case	Key gases	Concentration (PPM)	Diagnosis	Label
1	C_2H_2	> 20	Discharges	1
		≤ 20	No Discharges	0
		No record	Undefined	NaN *
2	H_2	> 100	Partial discharges	1
		≤ 100	No partial discharges	0
		No record	Undefined	NaN *
3	ΣC_xH_y (C_1,C_2)	> 500	Thermal fault	1
		≤ 500	No thermal fault	0
		No record in any gas	Undefined	NaN *
Diagnosis: [Case 1 | Case 2 | Case 3] (Diagnosis is in vector form, label is 1 if a fault is indicated and 0 if no fault is indicated)				
Normal condition for key-gas concentrations: 0 | 0 | 0				

* NaN: Not-a-Number.

Table 4.12: Fault-diagnosis table for CIGRÉ's key-gas ratios

Case	Key gases	Limit	Diagnosis	Label
1	$\frac{C_2H_2}{C_2H_6}$	> 1	Discharges	1
		≤ 1	No Discharges	0
		No record in one or both gases	Undefined	NaN *
2	$\frac{H_2}{CH_4}$	> 10	Partial discharges	1
		≤ 10	No partial discharges	0
		No record in one or both gases	Undefined	NaN *
3	$\frac{C_2H_4}{C_2H_6}$	> 1	Thermal fault	1
		≤ 1	No thermal fault	0
		No record in one or both gases	Undefined	NaN *
Diagnosis: [Case 1 | Case 2 | Case 3] (Diagnosis is in vector form, label is 1 if a fault is indicated and 0 if no fault is indicated)				
No interpretation for key-gas ratios: 0 | 0 | 0				

* NaN: Not-a-Number.

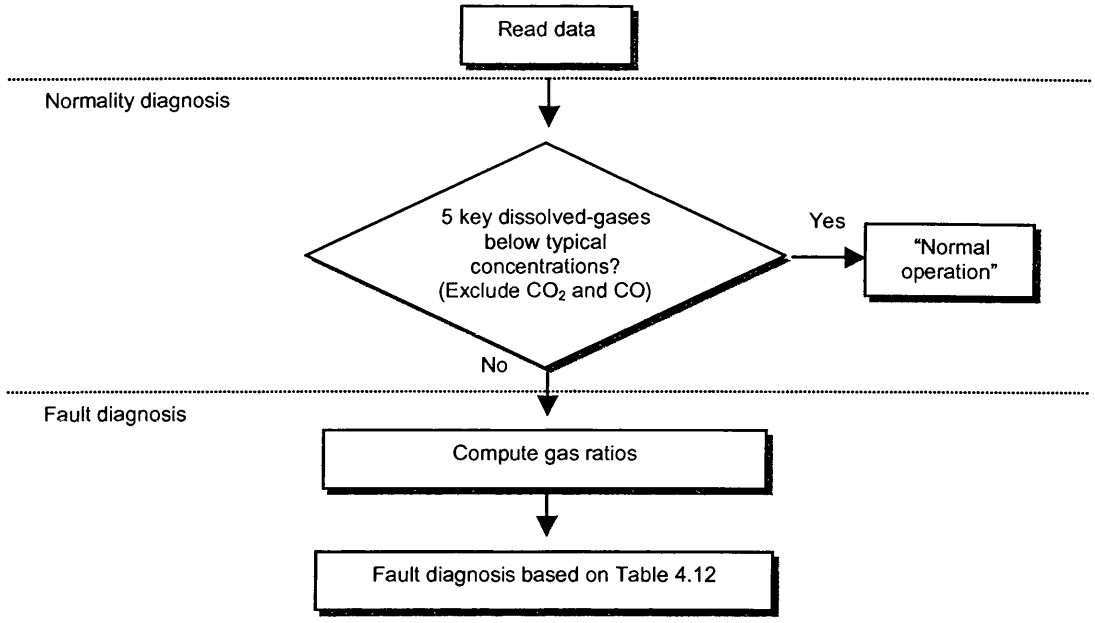


Figure 4.6: Flow-chart for the implementation of CIGRÉ's key-gas ratios

4.9 Comparative Study on Implemented DGA Schemes

The capability of the foregoing implemented schemes for the interpretation of DGA data was compared based on a set of 755 actual DGA records, as shown in Table 4.13. This set of DGA data was constructed from six subsets of actual DGA records from power transformers of three manufacturers, two voltage levels and one power rating.

Table 4.13: DGA data for comparative study on the implemented DGA schemes

Subset	Manufacturer *	Voltage Level (kV)	Power Rating (MVA)	Number of Records
1	I	400/132	240	72
2	I	275/132	240	91
3	II	275/132	240	187
4	II	400/132	240	94
5	III	400/132	240	108
6	III	275/132	240	203

* Roman numerals are used to represent different manufacturers.

Since each DGA interpretation scheme has its own interpretation of transformer condition, it is difficult to conduct the comparison unless only general conditions are considered, i.e. normal operation, discharges, PD and TF. Hence, the implemented DGA schemes were employed for the interpretation of actual DGA data (see Table 4.13) with the aim of detecting four general conditions as mentioned above. Table 4.14 illustrates the outcome of the comparison. Note that the number within each cell of Table 4.14 represents the amount of actual DGA records that were diagnosed to indicate a certain condition by using each of the implemented DGA schemes.

Table 4.14: Interpretation of actual DGA data using implemented DGA schemes

Diagnosis / DGA Schemes	N [a]	D [b]	PD [c]	TF [d]	EF [e] (D+PD)	D + TF	PD + TF	EF + TF	N/A [f]	N/I [g]
Dörnenburg Ratios	446	89	0	45					151	24
Rogers Ratios	484	156	0	47						68
IEC-1978 Ratios	484	167	0	71						33
IEC-1999 Ratios	484	157	0	75						39
Duval Triangle (1980 edition)	484	180	0	91						
Duval Triangle (1993 edition)	484	180	0	82		9				
CIGRE Key-Gas Concentrations	522	53	3	21	84	2	15	55		
CIGRE Key-Gas Ratios	484	2	0	68	0	193	1	0		7

a. N: Normal operation
b. D: Discharges
c. PD: Partial discharges
d. TF: Thermal fault
e. EF: Electrical fault
f. N/A: Not applicable (for Dörnenburg Ratios only)
g. N/I: No Interpretation
h. Areas marked by grey colour are not applicable to respective DGA interpretation schemes.

As can be observed from Table 4.14, the application of different DGA interpretation schemes on an identical set of DGA data have resulted in diverse interpretations of transformer condition, which could be potentially disrupting for those power utilities who wish to utilise these schemes for the condition assessment of their transformer population. In addition, some of the DGA schemes are not capable of interpreting some DGA records, as evident by the amount of DGA records in the "no interpretation" column of Table 4.14. Moreover, the application of CIGRÉ's key-gas concentrations and key-gas ratios could be confusing since a number of conditions were identified, which included simultaneous occurrences of two or three incipient faults.

It is apparent from the foregoing that an improved approach of DGA interpretation, which potentially does not give rise to the confusing circumstances as described previously, is needed in order to improve the confidence level and accuracy of the DGA interpretation. The proposed approach for achieving the foregoing objective will be described in subsequent chapters of the thesis.

4.10 Summary

The implementation of various DGA interpretation schemes has been presented in this chapter. The objective for implementing these schemes is to compare the strengths and weaknesses of various established schemes. In fact, it has been verified in Section 4.9 that interpretations as provided by these schemes are not compatible with one another, which could be potentially confusing for those power utilities who wish to apply these schemes for condition assessment of their power transformers. Thus, it is important to develop a novel approach that can significantly minimise the foregoing confusing situation, thereby improving the confidence level and accuracy of the DGA interpretation.

CHAPTER 5

KNOWLEDGE DISCOVERY METHODOLOGY FOR EXPLORATORY DATA ANALYSIS

5.1 Introduction

This chapter introduces a novel methodology for exploratory data analysis known as the knowledge discovery in databases (KDD), in which data mining (DM) is a vital component within the KDD methodology. This chapter will emphasis on a class of DM methods that are capable of performing clustering and classification functions while also able to present the "learned" information in an unambiguous and discernible manner.

Two relevant DM methods have been investigated in the research project, i.e. Sammon mapping [49] and the self-organising map (SOM) [50]. However, it was found that Sammon mapping offers a limited visualisation capability and it requires a very intensive computing resource; thus it is deemed not practical for analysing usually large and complex practical databases. Therefore, the SOM has been chosen as the suitable DM method worthy of further investigation in this research project due to its powerful visualisation capability, fast and efficient learning process, and comparatively light requirement on computing resource. Various issues for practical application of SOM will be presented in this chapter. The chapter concludes with a discussion on the feasibility of the SOM algorithm for exploratory data analysis on the condition assessment data of power transformers.

5.2 The Need for Novel Methodology of Data Analysis

The technology of computing and storage has enabled people to collect and store information from a wide range of sources at rates that were, only a few years ago, considered unimaginable. Although modern database technology enables economical storage of this data, we still do not have the technology to fully assist us analyse, understand or even visualise this stored data [51]. Consequently, the huge research interests in the afore-mentioned requirements have prompted the establishment of a new methodology known as the KDD, in which DM represents a key component of the KDD methodology.

Why are today's database and automated match and retrieval technologies not adequate for addressing these requirements? The answer lies in the fact that the patterns to be searched for, and the models to be extracted are typically subtle and require significant domain knowledge. In the past, we could rely on human analysts to perform the necessary analysis. Essentially, this meant transforming the problem into one of simply retrieving data, displaying it to an analyst, and relying on expert knowledge to reach a decision. However, with large databases, a simple query can simply return hundreds or thousands of matches; presenting the data, letting the analyst digest it, and enabling a quick and correct decision becomes unfeasible. Although data visualisation techniques can significantly assist this process, ultimately the reliance on humans in the loop becomes a major bottleneck.

Furthermore, there are also situations where one would like to search for patterns that humans are not well suited to find. Typically, this involves statistical modelling, followed by outlier detection, pattern recognition over large data sets, classification or clustering. Most database management systems do not allow the type of access and data manipulation that these tasks require; there are also serious computational and theoretical problems attached to performing data modelling in high-dimensional spaces and with large amounts of data.

The foregoing challenges are central to KDD and need urgent attention. Without heavily emphasising the development and research of KDD, we run the risk of forfeiting the value of most of the data that we collect and store. We would eventually drown in an ocean of massive (but valuable) data sets that are rendered useless because we cannot distil the essence from the bulk. To draw on the data-mining analogy, the precious nuggets of knowledge need to be extracted and the massive raw material needs to be managed appropriately and preferably recycled effectively [51].

5.3 Knowledge Discovery in Databases

The KDD can be defined as the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [51]. More specifically, KDD is a process of using DM methods or algorithms to extract or identify what is deemed knowledge according to the specifications of measures and thresholds, using the database along with any required pre-processing, sampling and transformation of the database [52]. Essentially, KDD is a process that is constituted of various interactive and iterative stages, as illustrated in Figure 5.1.

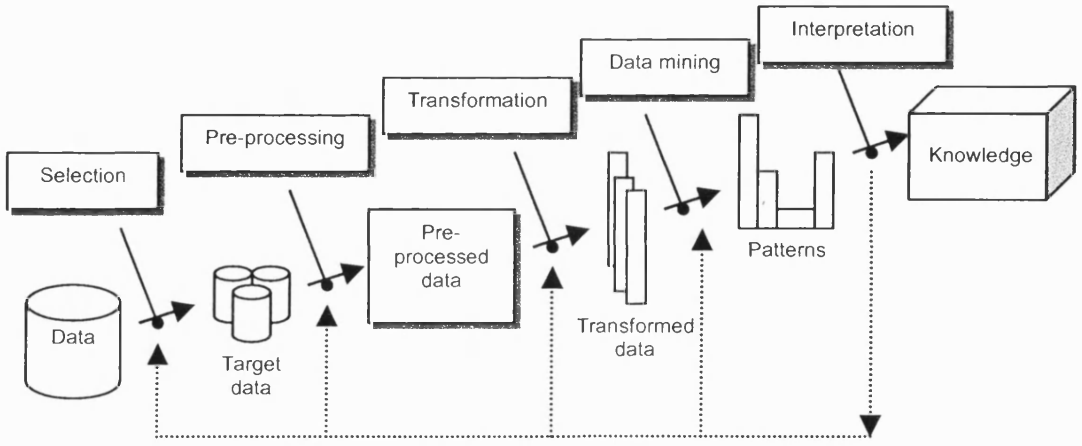


Figure 5.1: The KDD process

In essence, the KDD process can be expanded into the following stages with reference to Figure 5.1:

- Stage 1: Developing an understanding of the application domain, the relevant prior knowledge, and the goals of the end user.
- Stage 2: Creating a target data set, selecting a data set, or focusing on a subset of variables or data samples, on which discovery is to be performed.
- Stage 3: Data cleaning and pre-processing.
- Stage 4: Data reduction and transformation.
- Stage 5: Choosing the DM task.
- Stage 6: Choosing the DM algorithm(s).
- Stage 7: Performing DM based on chosen task and algorithm(s).
- Stage 8: Interpreting "mined" patterns, possible to return to any of Stage 1 to 7 for further iteration.
- Stage 9: Consolidating discovered knowledge.

The KDD process can involve a significant number of iterations and may contain loops between any two stages. Small changes at one stage can dramatically affect the rest, and consequently can make the difference between success and failure of the KDD process.

5.4 Data Mining: A Key Component of the KDD Process

DM is defined as a step in the KDD process consisting of particular algorithm(s) that, under some acceptable limitations of computational efficiencies, produces a particular enumeration of patterns over the database [52]. The DM process involves fitting models to, or determining patterns from, observed data. The fitted models play the role of inferred knowledge; whether or not the models reflect useful or interesting knowledge is part of the overall, interactive KDD process where subjective human judgement is usually required. Most DM methods are based on the statistical approach whereby uncertainties about the exact nature of real-world data generating

processes are taken into consideration throughout the modelling process. Specifically, most DM methods are based on concepts from various fields such as machine learning, pattern recognition and statistics. Various aspects of DM are briefly discussed in the following sections:

5.4.1 Primary Tasks of Data Mining

In practice, the two high-level primary goals of DM tend to be prediction and description. Prediction involves using some variables or fields in the database to predict unknown or future values of other variables of interest. On the other hand, description focuses on finding human-interpretable patterns describing the data. The relative importance of prediction and description for particular DM applications can vary considerably. However, in the context of KDD, description tends to be more important than prediction. The goals of prediction and description are achieved by using the following primary DM tasks [52]:

- Classification, in which a function is learned which maps or classifies a data item into one of several pre-defined classes.
- Regression, in which a function is learned which maps a real-valued prediction variable.
- Clustering, in which a common descriptive task is used to identify a finite set of categories or clusters to describe the data.
- Summarisation, in which a compact description is found for a subset of data.
- Dependency modelling, in which a model is found which describes significant dependency between variables.
- Change and deviation detection, in which significant changes in data are discovered from previously measured values.

5.4.2 Components of Data Mining Algorithms

Three primary components can be identified in any DM algorithm: model representation, model evaluation, and search [52]. This reductionist view is not necessarily complete or fully encompassing; rather it is a convenient way to express the key concepts of DM algorithms in a relatively unified and compact manner.

The model representation is the language for describing discoverable patterns. If the representation is too limited, then no amount of training time or examples will produce an accurate model for the data. Thus it is important to fully comprehend the representational assumptions that may be inherent to a particular method. It is also equally important that an algorithm designer clearly states which representation assumptions are being made by a particular algorithm. Note that more powerful representational power for models increases the danger of over-fitting the training data, which results in reduced prediction accuracy on unseen data. In addition, the search becomes much more complex and interpretation of the model is typically more difficult.

The model evaluation estimates how well a particular pattern meets the criteria of the KDD process. The evaluation of predictive accuracy is based on cross-validation while the evaluation of descriptive quality involves predictive accuracy, novelty, utility and understandability of the fitted model. Both logical and statistical criteria can be used for model evaluation.

The search method consists of two components: parameter search and model search. In parameter search, the algorithm must search for the parameters that optimise the model evaluation criteria, given observed data and a fixed model representation. For relatively simple problems there is no search; the optimal parameter estimates can be obtained in closed form. Typically, for more general models, a closed form solution is not available; "greedy" iterative methods are commonly used. On the other hand, model search occurs as a loop over the parameter search method; the model representation is changed so that a family of models are considered. For each specific

model representation, the parameter search method is utilised to evaluate the quality of that particular model. The implementation of model search methods tends to use heuristic search techniques since the size of the space of possible models often prohibits exhaustive search and closed form solutions are not easily obtainable.

5.4.3 Data Mining Methods

There exist a wide variety of DM methods, for example statistical-based methods, neural networks (NNs), evolutionary programming, memory-based reasoning, decision trees, genetic algorithms (GAs) and non-linear regression methods. The DM method explored in this chapter, i.e. the SOM, belongs to the unsupervised-NNs category.

5.4.4 Commercial Data Mining Software

Commercial DM software tools are available to assist commercial and public sectors to better manage and utilise their data. A prime example of DM software is Clementine [53], which is a leading toolkit that employs various modelling algorithms such as NNs, SOMs and regression-based techniques for data mining applications. It has been utilised for customer segmentation/profiling for marketing companies, fraud detection, credit scoring, load forecasting for utility companies and profit prediction for retailers. Another well-known DM software is DataEngine [54], which is a collection of tools for intelligent data analysis utilising fuzzy technologies, neural networks and conventional statistics. It has been successfully applied to the field of forecasting, database marketing, quality control, process analysis and diagnosis. Other example of DM software is DataScope [55], which comprises of Explorer, Predictor, Clusterer and Decision Support modules capable of creating a powerful knowledge discovery and mining environment suitable for wide range of industries and applications. Lastly, IRIS Explorer [56] is a DM software that employs a broad range of visualisation techniques, from simple graphs to multidimensional animation, in order to assist user to discern trend and relationships in the data. Other examples of commercial DM tools are well documented in [57].

5.5 Self-Organising Map: A Data Mining Method

The SOM [50] is a neural network (NN) algorithm based on unsupervised learning in a data-driven manner. It is essentially a non-linear mapping algorithm which projects a higher dimensional input space onto a lower dimensional output space while preserving the topological relations of the input space. This is accomplished via the self-organising process of the reference vectors of SOM, which is principally based on the Euclidean distances between reference vectors and input vectors. If sufficient training is performed on the SOM, an optimum approximation of the probability density function (PDF) of the input space can be achieved via a set of reference vectors that are arranged in a lower dimensional space.

The SOM can thus be applied for "unearthing" the hidden structure or inherent characteristics of a highly complex or multi-dimensional data and displaying them in a discernible and comprehensible manner. Owing to the foregoing capability of SOM, it has been applied successfully to various fields such as knowledge extraction [58, 59], classification [60], monitoring and modelling of complex processes [61], analysis of pulp and paper industries [62] and correlation hunting [57].

5.5.1 The SOM Algorithm

The SOM is formed by neurons located on a regular one- or two-dimensional grid. Although a higher dimensional grid can also be utilised, it is best avoided due to the difficulty in visualisation. The neurons can be arranged in hexagonal or rectangular configuration; the former is preferred owing to its effectiveness in visualisation. Figure 5.2 illustrates a two-dimensional SOM whose neurons are arranged in a hexagonal lattice.

The SOM essentially defines a mapping from the input space, $\Re^n$, onto a two-dimensional array of neurons. Every neuron, i, is associated with an n-dimensional reference vector, m_i, where $m_i = [\mu_{i1}, \mu_{i2}, \mu_{i3}, \ldots, \mu_{in}]^T \in \Re^n$, where n is the dimension

of the input space (i.e. number of components within an input vector). Hence, an input vector, x, where $x = [\xi_1, \xi_2, \xi_3, \ldots, \xi_n]^T \in \Re^n$, is connected to all neurons in parallel via scalar values $\mu_{ij}, j = 1, \ldots, n$, which are different for every neuron.

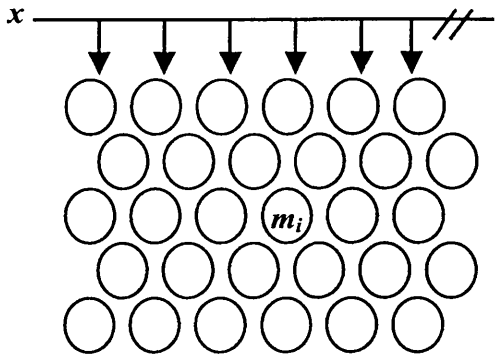
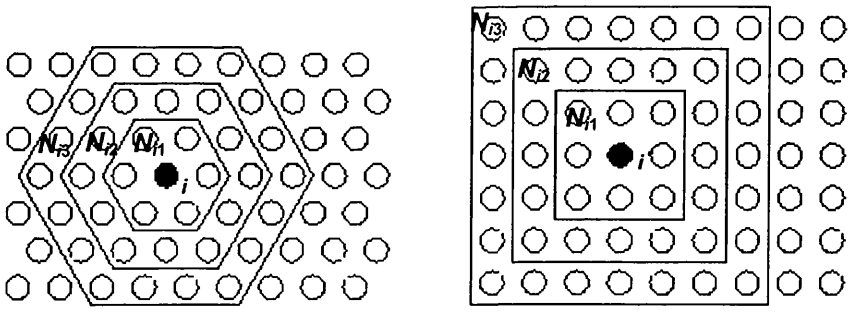


Figure 5.2: Two-dimensional SOM with neurons arranged in a hexagonal lattice

In addition, each neuron is associated with adjacent neurons of the SOM by a neighbourhood relation, which dictates the "area of influence" in the grid. Therefore, the immediate neighbours of neuron i, which are closest to neuron i, are associated with the first-neighbourhood, N_{i1}, of neuron i. The neighbourhood relations of different sizes in rectangular and hexagonal lattices are illustrated in Figure 5.3 [64].



(a) Hexagonal lattice **(b) Rectangular lattice**

Figure 5.3: The neighbourhood relation of SOM

The training process of SOM comprises of three stages, as explained in the following sections.

5.5.1.1 Initialisation of reference vectors

Reference vectors of SOM are initialised before the commencement of the training. This can be accomplished using three distinct approaches [50]. Generally, a random approach can be adopted, in which the reference vectors are initialised using arbitrary values; all initially unordered reference vectors will become organised after a few hundred training time-steps. An alternative strategy is to initialise the reference vectors based on randomly selected vectors from the input space, which is known as the sample initialisation. This is done so as to have a rough approximation of the density function of the input space before the start of the training process. Lastly, another approach of initialisation is to first determine the two principal eigenvectors of the autocorrelation matrix of the input data; reference vectors are initialised in an orderly fashion along the linear subspace spanned by the two principal eigenvectors. The foregoing approach is known as the linear initialisation. The important point to note is that the SOM is robust with regard to any initialisation approach. If the initialisation process is properly conducted, using one of three foregoing approaches, the SOM will converge faster to an optimum solution.

5.5.1.2 The training process

Once reference vectors of SOM have been initialised, the training of SOM commences by first choosing an input vector, x, randomly from the input data at time-step t. Comparison is then performed, with the computation of similarity measures, between x and all reference vectors of SOM. The best match of x on the SOM grid is known as the best matching unit (BMU), denoted as c, which has the maximum similarity (or nearest distance) with respect to x. The neuron c is thereby the location of the "response" of SOM towards the input vector, x. In simple terms, the input vector, x, is "mapped" onto the SOM at location c. The similarity is usually

defined by a distance measure, usually Euclidean distance. The foregoing process is defined by Equation (5.1):

$$c = \arg\min_i \{\|x - m_i\|\} \text{ or } \|x - m_c\| = \min_i \{\|x - m_i\|\} \quad \text{.....(5.1)}$$

Once the BMU, c, has been determined, all reference vectors of SOM are updated. The updating process is performed in such a way that those neurons that are topologically close to the BMU, c, up to a certain geometrical distance will also be activated and "learned" from the same input vector, x. This updating process will "stretch" the BMU and its topological neighbours towards the input vector, x, as illustrated in Figure 5.4 [64].

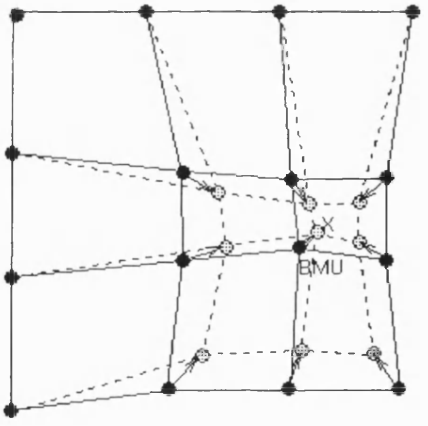


Figure 5.4: Influence of input vector towards its BMU and neighbours

A continuous random presentation of input vector, x, at each subsequent time-step, t, will result in a local "relaxation" or "smoothing" effects on the reference vectors of neurons in the neighbourhood. If viewed from the input space, these reference vectors will eventually form an "elastic" network that twists and folds onto the "cloud" of points in the input space that are formed by input vectors. The foregoing description will be graphically illustrated in Section 5.5.3.

The updating process of reference vectors is defined by Equation (5.2), as shown below.

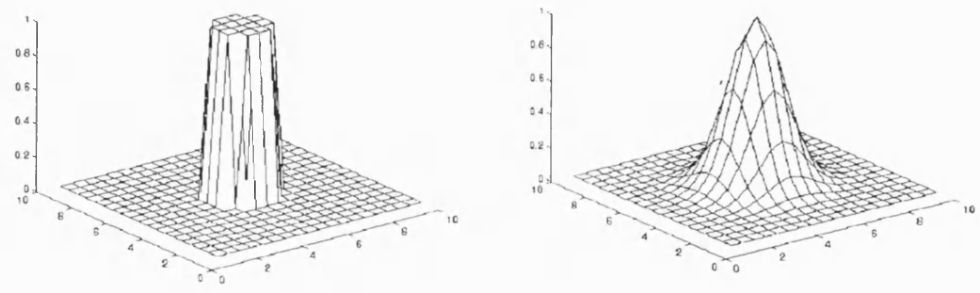
$$m_i(t+1) = m_i(t) + h_{ci}(t)\left[x(t) - m_i(t)\right] \quad \text{.....(5.2)}$$

Where, $h_{ci}(t)$ is the neighbourhood kernel around the BMU, c, at time-step, t. It plays a very important role in the self-organising process of reference vectors. Essentially, it acts as a smoothing kernel that is defined over the BMU, c, and its neighbours, at time-step, t. The neighbourhood kernel, $h_{ci}(t)$, is defined as:

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c - r_i\|, t) \quad \text{.....(5.3)}$$

Where, $\alpha(t)$ $(0 < \alpha(t) < 1)$ is the learning-rate function and $h(\|r_c - r_i\|, t)$ is known as the neighbourhood function. $r_c \in \Re^2$ and $r_i \in \Re^2$ are the location vectors of neurons c and i respectively. For convergence, it is necessary that $h_{ci}(t) \rightarrow 0$ when $t \rightarrow \infty$, and, with increasing $\|r_c - r_i\|$, $h_{ci} = 0$.

There are two common choices for the neighbourhood function: rectangular function and gaussian function, as illustrated in Figure 5.5 [64].



(a) Rectangular function **(b) Gaussian function**

Figure 5.5: Neighbourhood functions of SOM

For rectangular function, the neighbourhood kernel can be defined as:

$$h_{ci}(t) = \begin{cases} \alpha(t), \text{ if } i \in N_c(t) \\ 0, \text{ if } i \notin N_c(t) \end{cases} \quad \text{.....(5.4)}$$

For gaussian function, the neighbourhood kernel can be defined as:

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)} \right) \quad \text{.....(5.5)}$$

Where, $\alpha(t)$ defines the width of the neighbourhood kernel, which is analogous to N_c in Equation (5.4). Note that the learning-rate function, $\alpha(t)$, neighbourhood, $N_c(t)$, and width of the gaussian function, $\sigma(t)$, are all monotonically decreasing functions of time-step, t; the exact form of these functions is found to have no significant impact towards the learning efficiency of SOM [50].

In summary, the training process of SOM comprises of following steps:

- Step 1: Initialise reference vectors of all neurons at time-step, $t = 0$.
- Step 2: At each time-step, t, select one input vector, $x(t)$, randomly from the input data.
- Step 3: At each time-step, t, determine the BMU, c, for input vector, $x(t)$ according to Equation (5.1).
- Step 4: Update all reference vectors according to Equation (5.2).
- Step 5: Check if user-defined iterations and criteria have been reached. If YES, then stop training; if NO, increase time-step, t, by 1 and repeat the process from step 2.

Various issues have to be considered for practical application of SOM; these will be discussed in Section 5.5.2.

5.5.1.3 Calibration of trained SOM

When a sufficient number of input vectors, $x(t)$, have been presented and all reference vectors, $m_i(t)$, have converged to reasonably satisfactory values in accordance with user-defined iterations and criteria, the next possible step is the calibration of the trained SOM. However, this calibration process can only be performed if the characteristics of input data or the inherent process of the generation of data is of a known entity. This task is accomplished by inputting a number of input vectors, which are associated with some known characteristics or features, to the trained SOM and looking at where the best matches actually lie; the SOM neurons can therefore be labelled correspondently. In the event that the characteristics of the input data are of unknown or ambiguous entity, the interpretation of trained SOM may have to be conducted via other means, for example through comparison or validation with other analogous approaches.

5.5.2 Practical Advices for Effective Application of SOM

Although the SOM algorithm is theoretically well defined, there are several issues that need to be addressed for the practical application of SOM to solve some real-world problems such as the analysis of highly complex data or processes. In fact, some of the settings for the SOM training have to be determined empirically or based on experiences owing to the lack of mathematically well-defined solutions for determining these settings. Various issues for the application of SOM are discussed in the following sections.

5.5.2.1 The configuration of SOM

There are three aspects that need to be considered in order to obtain a good configuration of SOM before the commencement of training: shape and dimension, total number of neurons, and the number of neurons in each dimension of SOM.

Firstly, a two-dimensional SOM in a rectangular shape is generally adequate for most problems. In addition, the neurons are arranged in hexagonal configuration for ease of visualisation.

Secondly, the total number of neurons required can be determined empirically via the following formula [65]:

$$num\_neuron = 5 \times \sqrt{dlen} \quad \text{.....(5.6)}$$

Where, *num_neuron* is the total number of neurons and *dlen* is the total number of input vectors. However, it was found that 4 × *num_neuron* is generally more appropriate due to the following consideration on the size of SOM for effective visualisation of the training outcome.

Thirdly, given a fixed total number of neurons in a two-dimensional, rectangular SOM, the number of neurons in the x- and y-dimension of the SOM (i.e. map-sizes) can be determined by calculating the two principal eigenvalues of the covariance matrix of input data [65]. The covariance matrix, C, of the input data, D, is calculated according to Equation (5.7).

$$C = COV(D, D) = E[(D - \mu_D)(D - \mu_D)] = \frac{1}{dlen}\sum_{i=1}^{dlen}(D_i - \mu_D)^2 \quad \text{.....(5.7)}$$

Where, D_i is the i^{th} input vector and μ_D is the vector of mean-values for components of the input data. If, for example, the first and second eigenvalues of the covariance matrix, C, are denoted as α and β respectively, the ratio of map-sizes for both x- and y-dimension can be determined according to the following equation [65]:

$$\frac{x - dim}{y - dim} = \sqrt{\frac{\alpha}{\beta}} \quad \text{.....(5.8)}$$

Where, *x-dim* and *y-dim* are the respective map-sizes (i.e. the number of neurons) for *x*- and *y*-dimension of the SOM. Actual values for *x-dim* and *y-dim* are then determined in such a way that the product of *x-dim* and *y-dim* is equal to *num_neuron*, while complying with the ratio as calculated from Equation (5.8). Hence:

$$(x\text{-}dim)\times(y\text{-}dim)=num\_neuron \quad \text{.....(5.9)}$$

Nevertheless, it is well established empirically that better visualisation of the training outcome can be obtained if *x-dim* and *y-dim* are increased by two-fold. Consequently, the calculated *num_neuron* from Equation (5.6) is increased by four-fold in order to cater for the increase in map-sizes.

Hence, the relationship between map-sizes and the total number of neurons are as shown in Equation (5.10). Note that the ratio of *x-dim* to *y-dim*, as calculated from Equation (5.8), must always be complied.

$$2(x\text{-}dim)\times 2(y\text{-}dim)=4\times num\_neuron \quad \text{.....(5.10)}$$

In general, the mapping efficiency of SOM does not suffer considerably in larger map-sizes. However, as the size of SOM increases, longer training intervals may be required so as to obtain a satisfactory outcome.

5.5.2.2 Determination of training parameters

Once a good configuration of SOM has been determined according to Section 5.5.2.1, the next-step is to consider some vital issues for the training of SOM:

- Selection of neighbourhood function.
- Selection of training method: sequential or batch?
- Determination of training phases.

- Selection of radius for each phase of training.
- Selection of learning-rate function for each phase of training.

Firstly, the two commonly used neighbourhood functions are the rectangular function and the gaussian function; other functions can also be employed. In essence, the choice of neighbourhood function depends on the objective of the end-user. If the SOM is used for "mining" high-dimensional or complex data, the rectangular function is found to be more suitable through experience due to the clear "separation" of inherent characteristics resulting from the trained SOM. Nonetheless, the gaussian function can also be applied, but the "separation" of inherent characteristics is found to be less apparent when compared with that of rectangular function. Thus, the rectangular function is chosen as the suitable neighbourhood function for the research project as reported in this thesis.

Secondly, there are two main variants of SOM with regard to the training approach: sequential training and batch training. The sequential mode is a traditional and original way of training, in which reference vectors are updated at each presentation of input vector. On the contrary, the batch mode of training is a faster option since all reference vectors are only updated after the presentation of all input vectors (i.e. after each epoch). Again, the choice of sequential or batch mode of training is dependent on the end-user. Nevertheless, the sequential training will take longer time when compared with the batch mode. The sequential mode of training is employed for the research project reported in this thesis.

Thirdly, the training of SOM can be performed in two phases: "rough ordering" and "fine tuning" [64]. However, the former phase is not needed if linear initialisation is adopted for assigning preliminary values for the reference vectors. Generally, the "rough ordering" phase should be commenced with a high learning-rate, large radius and relatively fewer training iterations. The objective of performing the foregoing phase is to vaguely organise the SOM neurons and corresponding reference vectors into a structure which approximately displays the inherent characteristics of the input data. After that, the "fine-tuning" phase is commenced with low learning rate, small

radius and higher number of training iterations. The foregoing phase is performed so as to fine-tune the vague "structure" or inherent characteristics of data into a clearer, more organised and comprehensible patterns.

Fourthly, the neighbourhood radii for "rough ordering" and "fine tuning" phases have to be determined before the commencement of training. Normally, a large radius is adopted for the "rough ordering" phase with the intention that all neurons can learn from the input vector, one way or other, at each training time-step, and their reference vectors are then adjusted accordingly. Heuristically, the neighbourhood radius for the "rough ordering" phase is taken from the maximum between *x-dim* and *y-dim*. In the "fine tuning" phase, however, a smaller radius, usually at a fraction of the radius adopted in the "rough ordering" phase, is utilised due to the fact that activation of all neurons is not necessary for learning since a rough data "structure" or inherent characteristics have already been obtained from the previous phase. Again, it is practically found that a fraction of 1/8 is generally suitable. Note that all radii will decrease monotonically to one at the end of each phase of training [65], in accordance with a simple function of training time-step, t.

Finally, the learning-rate for the "rough ordering" phase is usually taken to be approximately ten times larger than that of "fine-tuning" phase. In fact, initial learning-rates of 0.05 and 0.001 are actually adopted in this research project for the "rough ordering" phase and the "fine tuning" phase, respectively. In practice, other values can also be adopted since the learning-rates generally only impact on the speed of training. However, it is beneficial to opt for smaller learning-rates since the updating process of reference vectors can be performed more efficiently in this way, and hence the inherent characteristics of input data can be learned more effectively. The learning-rate is normally a monotonically decreasing function of training time-step, t; learning-rates for both training phases will decrease from the foregoing initial values to about 1/100 of initial values at the end of both training phases [65].

5.5.2.3 Visualisation of the trained SOM

In essence, the trained SOM is represented by a collection of reference vectors, which have been adjusted in accordance with a set of input vectors. If the input data is of a simple two-dimensional or three-dimensional format (i.e. two or three components in an input vector), then the trained SOM can be easily "visualised" via the illustration of two- or three-dimensional plots of the reference vectors, since the dimension of reference vectors is the same as the dimension of input data (see Section 5.5.1). Nonetheless, if the input data is more than three-dimensional by nature, it is virtually impossible or mathematically very complex to "visualise" the trained SOM using the foregoing approach.

An excellent approach of visualisation has been proposed by Ultsch [58-60], which is known as the "unified-matrix" (u-matrix) method. The u-matrix method is essentially based on the Euclidean distances between SOM neurons that are located on the SOM grid. Essentially, it allows the visualisation of multi-dimensional reference vectors in a two-dimensional format, which is discernible and comprehensible to the end user. It effectively allows the visualisation of hidden "structure" or inherent characteristics of the input data, which have been "learned" or "mapped" by the trained SOM. Fundamental principles of the u-matrix method are presented as follow.

For a SOM that comprises of 5 neurons, in which each neuron is represented by a reference vector, $m(i)$:

$m(1)$; $m(2)$; $m(3)$; $m(4)$; $m(5)$

the u-matrix is a 9×1 vector which consists of the following components:

$u(1)$; $u(1,2)$;$u(2)$; $u(2,3)$; $u(3)$; $u(3,4)$; $u(4)$; $u(4,5)$; $u(5)$

Where, $u(i, j)$ is the Euclidean distance between the reference vector $m(i)$ and $m(j)$.

The $u(k)$ is calculated from $u(i, k)$ and $u(k, j)$ based on a user-defined function, such as the mean, maximum, minimum or median function, as defined in Equation (5.11).

$$u(k) = f(u(i,k), u(k,j)) \quad (5.11)$$

Thus, if the Euclidean distances, $u(i, j)$, among a group of SOM neurons are small, the resulting $u(k)$ will be small as well. Consequently, if these Euclidean distances are colour-coded by using a colour or grey-scale, the resulting similarity or dissimilarity among SOM neurons can thereby be identified via the means of a colour or grey border around a group of neurons of similar characteristics. Hence, the u-matrix can be employed for visualisation and summarisation of the hidden structure or inherent characteristics of the input data. In addition, inherent characteristics can also be displayed via the visualisation of component-planes in which trained values for each component of the input data are displayed, again via colour-coding by utilising either colour- or grey-scale. Examples on the visualisation of SOM will be demonstrated in Section 5.5.3.

5.5.2.4 The selection of optimum SOM

Given a SOM with a pre-determined configuration and pre-defined training parameters, how do we identify the number of training iterations needed in order to produce an optimum SOM for a set of input data? The answer is not so straight-forward due to the fact that no "targets" are available for the validation of SOM during the training process, similar to other types of unsupervised-NNs. Therefore, a combination of mathematical and heuristic approaches must be adopted so as to identify the optimum SOM for a given set of input data in a systematic and unambiguous manner.

Generally, there are two contradictory objectives that are involved in the training of SOM [64]. The first objective of applying SOM is for the purpose of "vector quantisation", i.e. to approximate a multi-dimensional input data using a set of

reference vectors which are arranged in a two-dimensional format (i.e. the SOM grid). For such a case, the Euclidean distance between each input vector and its BMU should be as close as possible so as to ensure that the trained SOM is the "best" approximation of the input data. On the contrary, it is also the objective of applying SOM for the visualisation of hidden structure or inherent characteristics of the input data. Thus, the trained SOM should also have the ability to generalise and not over-fit the training data.

Unfortunately, there is no single parameter which seamlessly measures the effectiveness of SOM for meeting the foregoing objectives. However, two SOM quality measures are available to calculate the ability of SOM for vector quantisation and generalisation respectively [65]. The ability of SOM to approximate a given set of multi-dimensional input data is measured by the "average quantisation error" (AQE). It is defined as the average Euclidean distance between input vectors and their BMUs. The AQE will tend to decrease with the increase in training iterations, since a longer training interval may lead to a gradual over-fitting of the training data.

The ability of SOM to generalise and preserve its topology is measured by the "topographic error" (TE). It is defined as the proportion of all input vectors for which their first and second BMUs are not adjacent to each other. The TE will tend to rise and fall at several stages throughout the training process, which are observed to be corresponding to the clarity of the hidden "structure" of the input data. However, a low TE does not necessarily means an optimum SOM since the corresponding AQE might be still quite high.

In view of the foregoing, a heuristic approach has been devised herein which apply the aforementioned SOM quality measures for selecting an optimum SOM for a particular set of input data. Firstly, a group of identically configured SOMs are trained using the same training parameters but only differ in the number of training iterations, which are increased gradually by a fixed amount. The training process is restarted for every SOM, i.e. after the initialisation of reference vectors for the "rough ordering" phase and after the "rough ordering" phase for the "fine tuning"

phase. Secondly, both AQE and TE are calculated after the training is completed for every SOM; a likely scenario for the magnitude of these measures is illustrated in Figure 5.6. Thirdly, several suitable candidates for the optimum SOM are selected from several "valleys" of the TE curve, as pointed out in Figure 5.6. Finally, the optimum SOM is selected from these candidates based on human visual inspection.

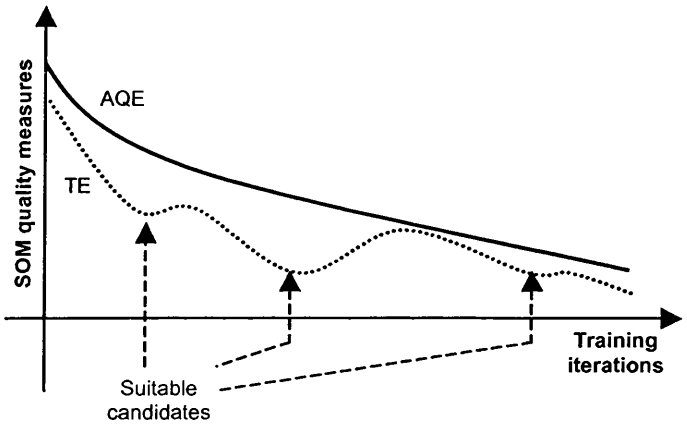


Figure 5.6: Selection of optimum SOMs based on several suitable candidates

5.5.3 Software Tools of SOM

Various software tools are available for those who are interested in using the SOM for data analysis. Among the SOM software tools are SOM_PAK [66] and its MATLAB implementation, the SOM Toolbox [67], which are developed by the Neural Network Research Centre (NNRC) at the Helsinki University of Technology, Finland. Since the NNRC is established by Prof. Teuvo Kohonen, the founder of the SOM algorithm, the above-mentioned software tools are the most faithful implementation of the SOM theories. Therefore, these software tools have been employed for conducting the research of data mining on transformer condition assessment data, as reported in subsequent chapters of this thesis. Particularly, the MS-DOS based SOM_PAK has been used for the SOM initialisation and training purposes and the MATLAB-based SOM Toolbox has been used for visualisation of trained SOM and post-processing on chosen optimum SOMs.

5.5.4 An Illustrative Example

The application of SOM for a simple data analysis is demonstrated. The input data contains 300 input vectors; there are 3 components in each input vector. The input data is generated by mixing three subsets, with each subset being constructed from 100 randomly generated vectors that are centred at [0 0 0], [2 1 1] and [1 2 1] respectively. Figure 5.7 illustrates the composition of the input data. The input data is plotted in Figure 5.8, with the centre of each subset being labelled as "+".

A SOM is configured in accordance with Section 5.5.2.1, which results in a hexagonal lattice of 24-by-14 neurons, as illustrated in Figure 5.9. In addition, training parameters are determined with reference to Section 5.5.2.2. Before the commencement of training, reference vectors of SOM are randomly initialised, resulting in the initial arrangement of SOM neurons being still "messy", as illustrated in Figure 5.10.

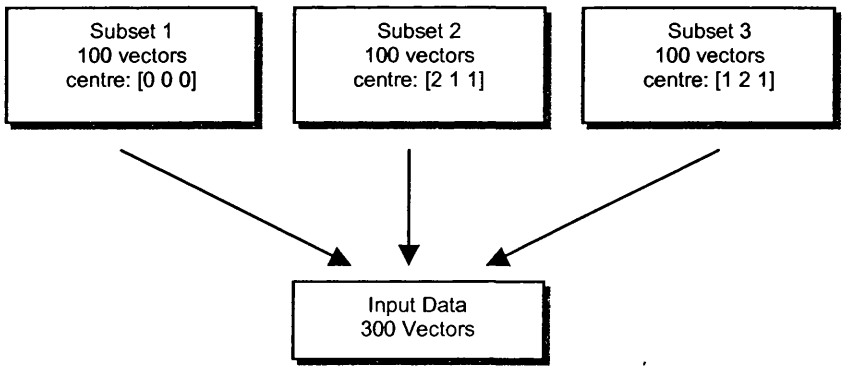


Figure 5.7: The composition of input data

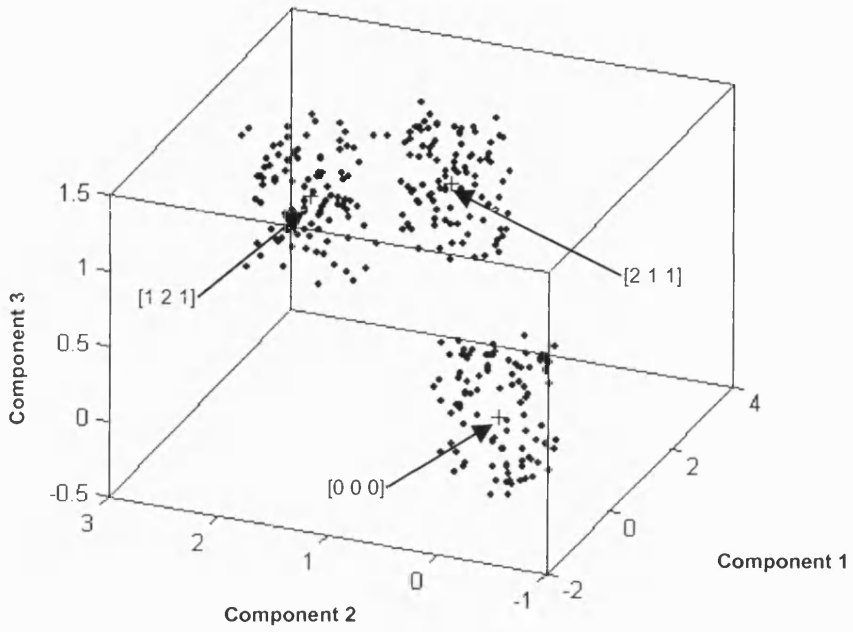


Figure 5.8: The input space with three subsets of the input data

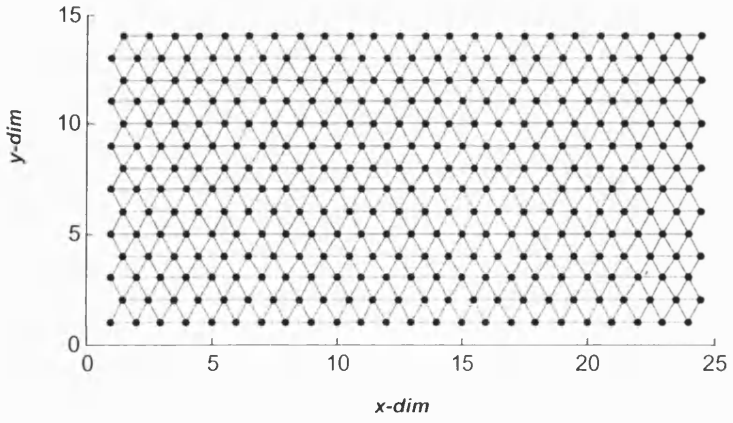


Figure 5.9: A SOM consists of 24-by-14 neurons arranged in hexagonal lattice

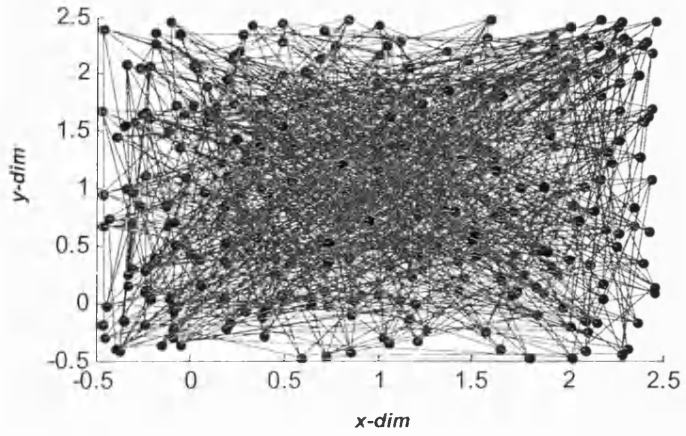


Figure 5.10: Arrangement of SOM neurons after the random initialisation

The training of SOM commences by first performing the "rough ordering" phase; an initial radius of 24 and an initial learning-rate of 0.05 are used for this phase. Figure 5.11 illustrates the arrangement of SOM neurons after the "rough ordering" phase, which takes approximately 3000 training time-steps to complete. The AQE is 0.3235 and the TE is 0.3067.

The reference vectors are further adjusted in the "fine tuning" phase, in which an initial radius of 3 and an initial learning-rate of 0.001 are used for this phase. Figure 5.12 illustrates the arrangement of SOM neurons in the input space after the "fine tuning" phase, which takes approximately 6300 training time-steps. The final AQE and TE are reduced to 0.3027 and 0.1600 respectively. As seen from Figure 5.12, the SOM now organises the input data in an excellent manner, in which shorter Euclidean distances between SOM neurons are found at the location of three subsets of input data.

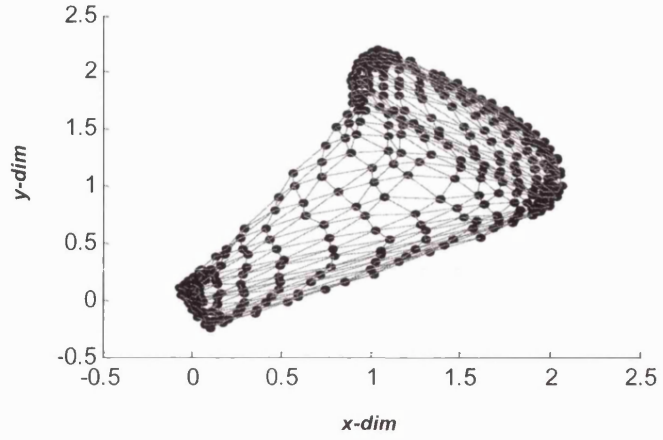


Figure 5.11: Arrangement of SOM neurons after the "rough ordering" phase

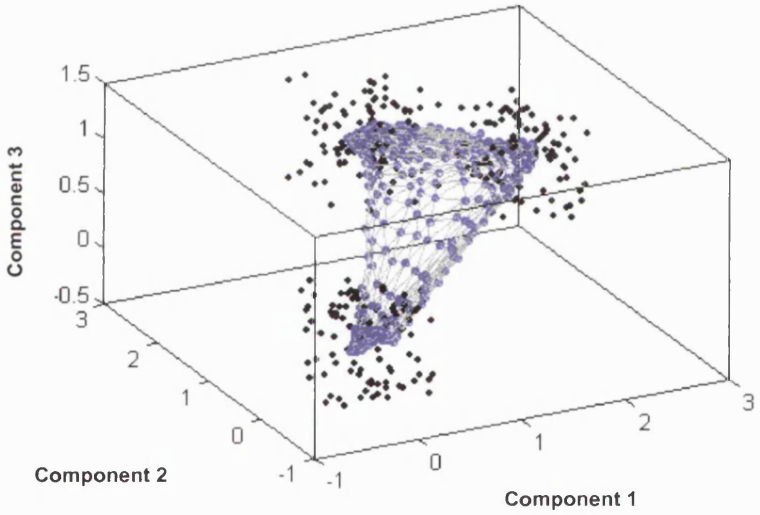


Figure 5.12: Arrangement of SOM neurons after the "fine tuning" phase

As described in Section 5.5.2.4, another approach of visualising the trained SOM is the visualisation of u-matrix and component planes, as illustrated in Figure 5.13. Colour bars on the right-hand-side of the u-matrix and component planes represent the magnitude, in which the colour bar on the u-matrix represents Euclidean distances between neurons and colour bars on component planes represent the magnitude of each component of the reference vectors.

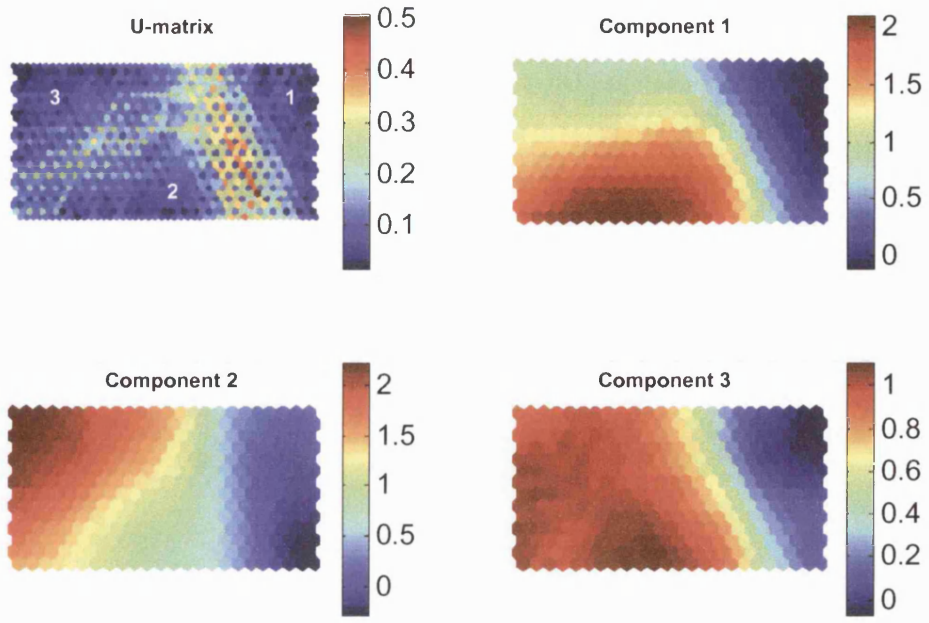


Figure 5.13: The u-matrix and component-planes of the optimum SOM

As observed from the u-matrix in Figure 5.13, the map is clearly divided into three regions of smaller Euclidean distances, in which each region actually corresponds to a subset of the input data. Therefore, the hidden "structure" or inherent characteristics of data can be observed through the use of the u-matrix visualisation.

In essence, the u-matrix and component-plane visualisation have to be employed for visualising the reference vectors of the trained SOM that are higher than three-dimensional by nature.

5.6 Feasibility of the DM Approach and the SOM for Exploratory Analysis on Condition Assessment Data

Dissolved gas analysis (DGA) is often performed on main units and peripheral devices (e.g. bushings and tap changers) of power transformers as part of the condition assessment procedures. Eventually, an enormous amount of DGA records

is gathered and stored in the database. These records actually contain lots of valuable information regarding the series of "events" or "incidences" (e.g. incipient faults) that have taken place in power transformers. If this information can be extracted, comprehended and visualised, our understanding of the health and condition of power transformers can be enhanced. The DM approach, particularly the SOM, can be employed to "mine" or "unearth" the hidden information or "knowledge" from the DGA data and display it in a user-understandable format. The feasibility of the SOM for the analysis and interpretation of DGA data will be investigated in subsequent chapters of the thesis.

Furthermore, the SOM can also be used for the analysis of sensor measurements so as to assess the inherent relationship between various sensor parameters and to summarise the operating condition of the power transformer. The feasibility of the SOM algorithm for performing these tasks will also be investigated in the thesis.

5.7 Summary

A novel methodology known as the KDD has been introduced in this chapter, in which the DM approach represents a vital component of the KDD process. A particular DM method known as the SOM has been described in detail. More importantly, various issues for the practical implementation of SOM have been thoroughly discussed and an illustrative example has been presented so as to demonstrate the capability of SOM for data analysis. Finally, the feasibility of the DM approach and the SOM in particular for the analysis of power transformer condition assessment data has been briefly discussed at the end of this chapter.

CHAPTER 6

STATISTICAL ANALYSES OF THE DGA DATA OF POWER TRANSFORMERS

6.1 Introduction

Statistical analyses were conducted on the dissolved gas analysis (DGA) data of power transformers, which was obtained courtesy of the National Grid Company (NGC), UK. The major aim of performing these statistical analyses is to gain an approximate insight into the statistical characteristics of the DGA data before it is subjected to a high-level and more sophisticated analysis by the proposed approach. Interesting features and characteristics were discovered throughout the analyses, which will be reported in subsequent sections of this chapter.

6.2 Format of DGA Data

The acquired DGA database comprises of 14943 DGA records dated from 1968 to 1998, which correspond to approximately 589 power transformers of the NGC plc., UK. The original format of the DGA data is illustrated in Table 6.1.

As shown in Table 6.1, nine dissolved gases and moisture (i.e. H_2O) are recorded in a single DGA record, accompanied by some background information of the corresponding power transformer from which the oil sample was drawn, such as its location (Plant_ID), identity (AMIS), manufacturer (Manc), date of commissioning (Comm), voltage level (KV), power rating (MVA), family (Fam), year of

manufacturing (Yrmd), date of sampling (Sampdate) and the percentage composition of the dissolved gases in an oil sample (Gas). Note that the unit of dissolved gases are in part-per-million (PPM), which represents a certain quantity of a dissolved gas, which is measured in micro-litre, in 1 litre of an oil sample. In addition, all oil samples were drawn from the bottom of transformer tanks, as indicated by the number "01" in the "Plant_ID".

Table 6.1: Format of the DGA records

A	B	C	D	E	F	G	H	I	J
Plant_ ID	**AMIS**	**Manc**	**Comm**	**KV**	**MVA**	**Fam**	**Yrmd**	**Samp date**	**Gas**
ABHA4 SGT1 01	66	EEC	1971	400/ 132	240	E07	1971	29/01/ 88	7.25
ABHA4 SGT1 01	66	EEC	1971	400/ 132	240	E07	1971	20/07/ 88	7.21
ABHA4 SGT1 01	66	EEC	1971	400/ 132	240	E07	1971	18/10/ 88	7.75

K	L	M	N	O	P	Q	R	S	T
CO	**CH_4**	**CO_2**	**C_2H_4**	**C_2H_6**	**C_2H_2**	**H_2**	**O_2**	**N_2**	**H_2O**
226	3	1460	2	1	1	15	0	0	9
252	3	2095	1	1	1	19	0	0	13
181	2	2340	1	1	1	18	0	0	9

6.3 Processing of DGA Data

The obtained DGA database was processed in order to eliminate several identified imperfections in the data, such as the presence of "blanks" (i.e. no data entries or records) and the ambiguity in the implication of "0 PPM" records for some dissolved gases. In addition, the format of which dissolved gases were recorded has to be standardised so as to ensure that similar statistical analyses can be applied to all dissolved gases.

Firstly, each "blank" was replaced by a common representation known as the "NaN" (i.e. Not-a-Number). The reason for keeping these blank entries is to preserve the valuable information contained within the original DGA data without interfering with its characteristics.

Secondly, some gases are always present in large quantities in the insulation oil if real-world circumstances are taken into consideration. Specifically, power transformers are not sealed-off completely from the surrounding environment and hence it is possible to have some atmospheric gases such as nitrogen (N_2), oxygen (O_2) and carbon dioxide (CO_2) dissolved in the insulation oil within the transformer tank. Moreover, the natural decomposition of cellulose insulation also produces CO_2 and carbon monoxide (CO); thus they are likely to present in large quantities in the insulation oil as well. In fact, records of "0 PPM" in these dissolved-gases could either be due to no measurement being taken for these gases or human errors. For that reason, the "0 PPM" records for the foregoing dissolved gases were replaced by "NaNs" in order to prevent confusion in the subsequent analyses.

Finally, a lot of measurements are recorded in the format of "$< x$", in which "x" is the "analytical detection limit" of each dissolved gas. This is due to the inability of the gas chromatograph (GC) to measure extremely low concentrations of some dissolved gases in the oil sample. Hence, these "$< x$" entries were converted to reasonably small values in a systematic manner in order to facilitate later statistical analyses, as shown in Table 6.2.

6.4 Categorisation of DGA Data

The processed DGA data was partitioned into the following data sets:

- DGA_KV: Partition of DGA data according to various voltage levels.
- DGA_MVA: Partition of DGA data according to various power ratings.

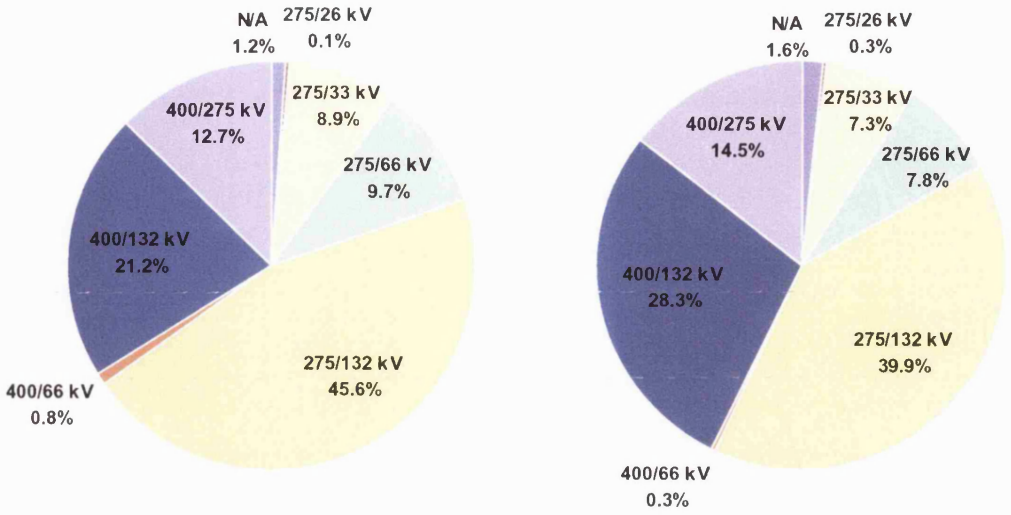
Table 6.2: Conversion of "< x" entries into assigned values

No.	Dissolved Gas	Original Format (PPM)	Assigned (PPM)
1	N_2	< 50	49
2	O_2	< 50	49
3	CO_2	< 50	49
4	CO	<10	9
5	H_2	< 1	0.9
6	CH_4	< 1	0.9
7	C_2H_6	< 1	0.9
8	C_2H_4	< 1	0.9
9	C_2H_2	< 0.2	0.19

The percentage composition of the DGA data according to various voltage levels and power ratings are illustrated in the form of pie charts, as illustrated in Figures 6.1 (a) and 6.3 (a). In addition, the number of corresponding power transformers were partitioned accordingly, as shown in Figures 6.1 (b) and 6.3 (b). Note that the notation of "N/A" in the foregoing pie charts signifies either unknown voltage levels or power ratings.

Some interesting characteristics can also be discovered when the amount of DGA data is plotted against the number of power transformers of the same voltage levels or power ratings, as shown in Figures 6.2 and 6.4 respectively.

First of all, it can be concluded that the relationship between the amount of DGA data and the number of power transformers is quite straightforward if data were to be divided according to various voltage levels, in which a larger number of power transformers of a certain voltage level, say 275/132 kV actually corresponds to the larger amount of DGA data. Moreover, a similar relationship is also observed if the amount of DGA data is plotted against the total age of power transformers for various voltage levels, as shown in Figure 6.5.



(b) Partition of DGA data according to various voltage levels of power transformers

(a) Partition of power transformers according to various voltage levels

Figure 6.1: Partition of DGA data and power transformers according to various voltage levels

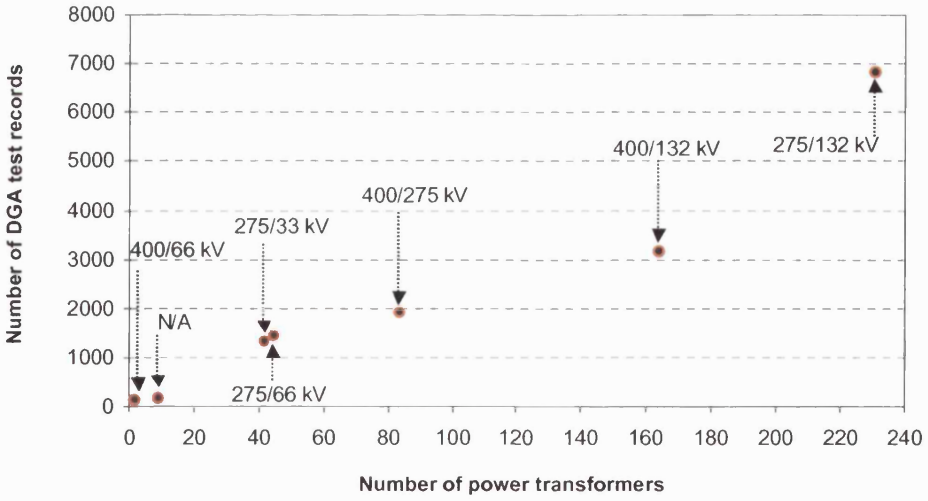
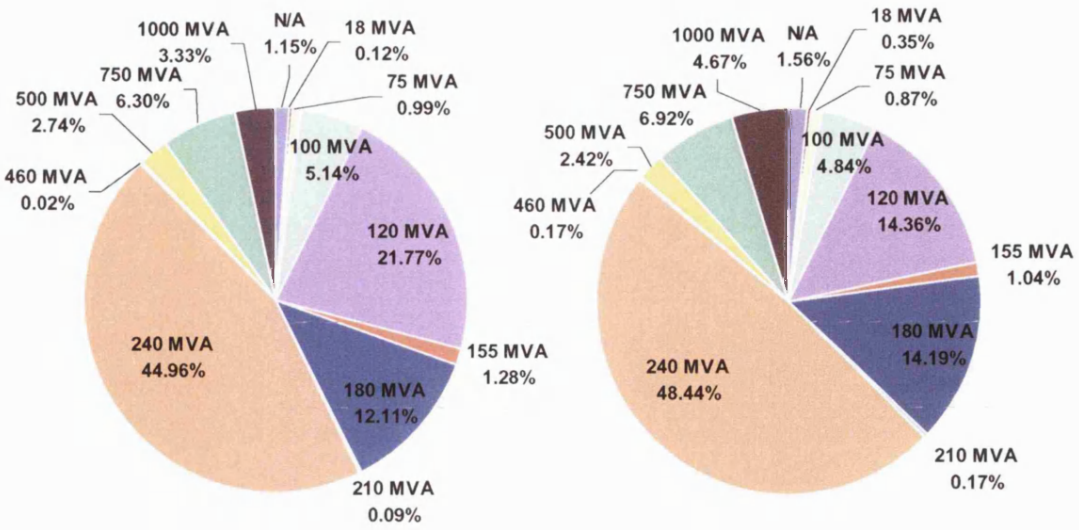


Figure 6.2: The amount of DGA data versus the number of power transformers according to various voltage levels



(a) Partition of DGA data according to various power ratings of power transformers

(b) Partition of power transformers according to various power ratings

Figure 6.3: Partition of DGA data and power transformers according to various power ratings

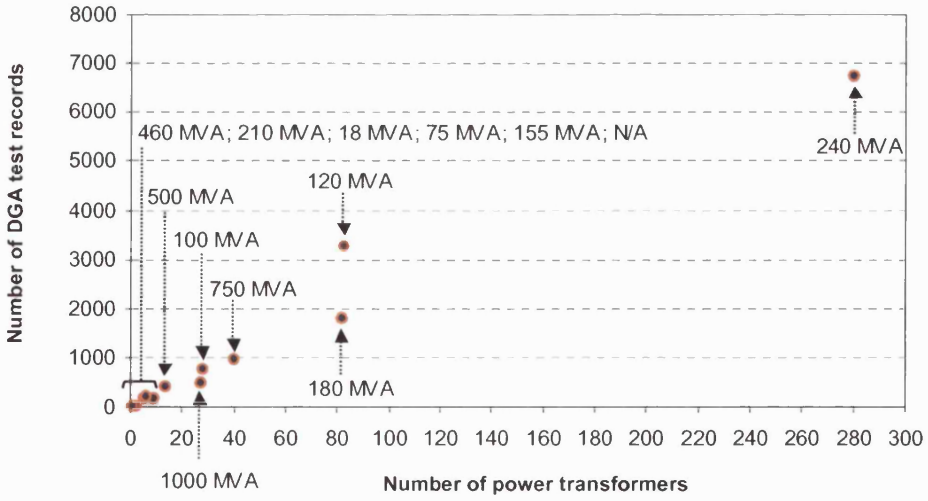


Figure 6.4: The amount of DGA data versus the number of power transformers according to various power ratings

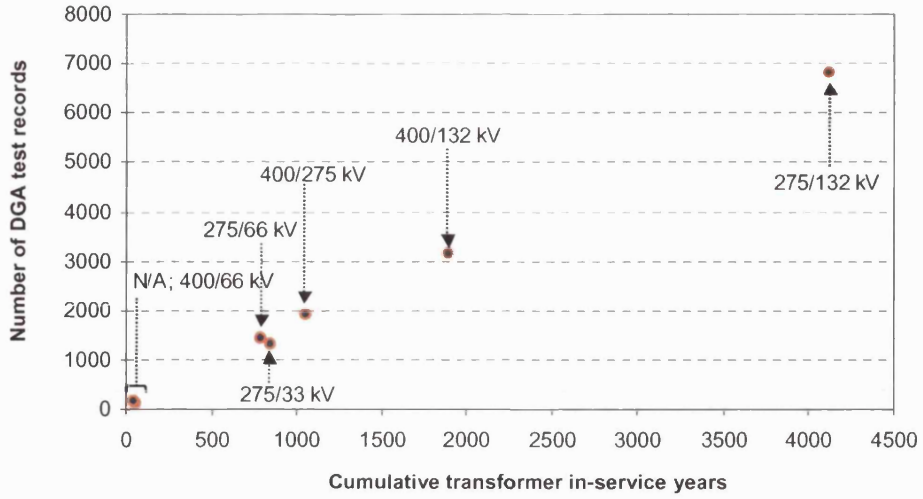


Figure 6.5: The amount of DGA data versus the total age of power transformers according to various voltage levels

Nevertheless, the relationship between the amount of DGA data and the number of power transformers is not so straightforward if data were to be divided according to various power ratings. As observed from Figure 6.4, although the number of power transformers that correspond to the power rating of 120 MVA and 180 MVA are quite similar, there is a significant difference in the respective amounts of DGA data. The reason for the foregoing becomes apparent when the amount of DGA data is plotted against the total age of power transformers for various power ratings, as illustrated in Figure 6.6. It is found that those power transformers within the category of 120 MVA are generally "older" if compared to those of 180 MVA, thereby forming a larger proportion of the DGA data.

In essence, there are various factors which determine the amount of DGA data contributed by a power transformer, such as its age, operating condition and frequency of sampling etc. The effect of age towards the amount of DGA data has been illustrated in Figures 6.5 and 6.6. Again, the notation of "N/A" in the foregoing figures means either unknown voltage levels or power ratings.

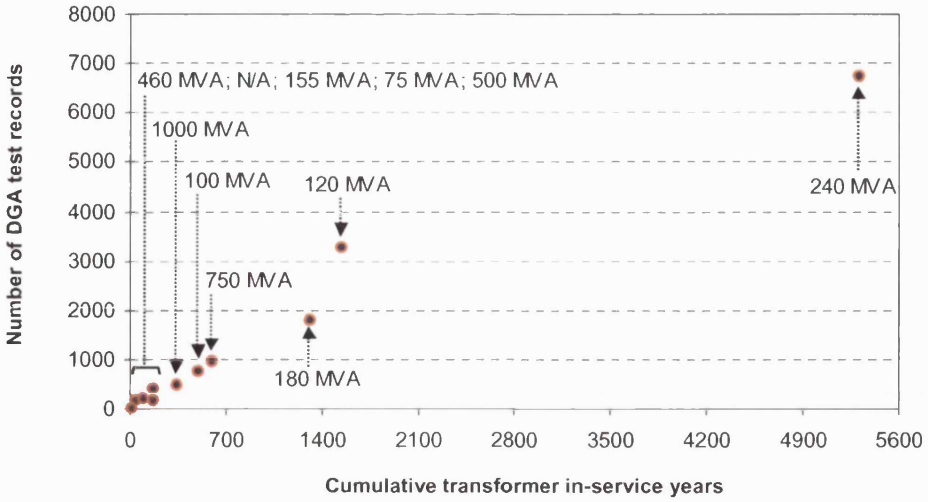


Figure 6.6: The amount of DGA data versus the total age of power transformers according to various power ratings

6.5 Frequency and Percentage Distribution of Dissolved Gases

The frequency distribution plot is also known as the histogram, which provides an excellent visualisation of the statistical distribution of dissolved gases. Basically, it shows the amount of data (i.e. frequency) that falls within a particular range of magnitudes. Besides, the percentage distribution was computed for each dissolved gas, by calculating the proportion of data that fall within a particular range of magnitudes out of the total amount of data. Finally, the frequency of "blanks" (i.e. no data entries) is shown in the negative-half of the x-axis; this is done without affecting the original characteristic of each dissolved gas.

Figures 6.7 to 6.15 illustrate the frequency and percentage distribution of dissolved gases; these plots have been zoomed-in so that only those regions of which the distribution of dissolved gases are mostly concentrated, are illustrated. As can be observed from these figures, at least 84% and above of the dissolved-gases are observed to be located within the lower range of concentrations relative to their maximum concentrations. There are only a few cases of which extremely high concentrations are observed for the dissolved gases.

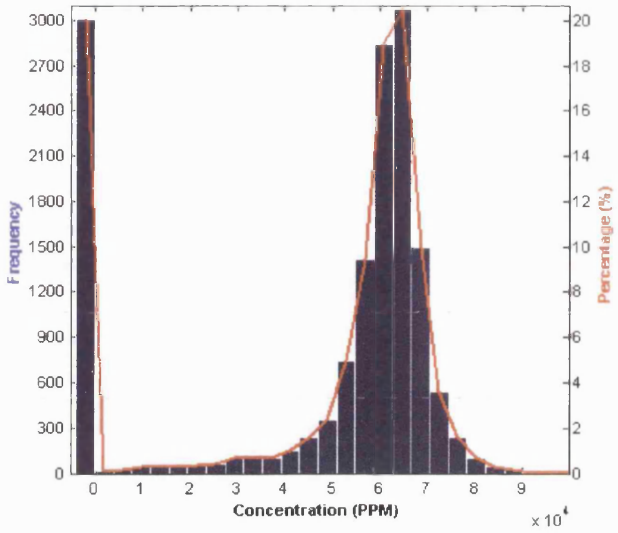


Figure 6.7: Frequency and percentage distribution of N_2

Note:
1. Number of blanks = 3008
2. Percentage data within the illustrated region = 99.81%
3. Maximum concentration of N_2 = 591766 PPM

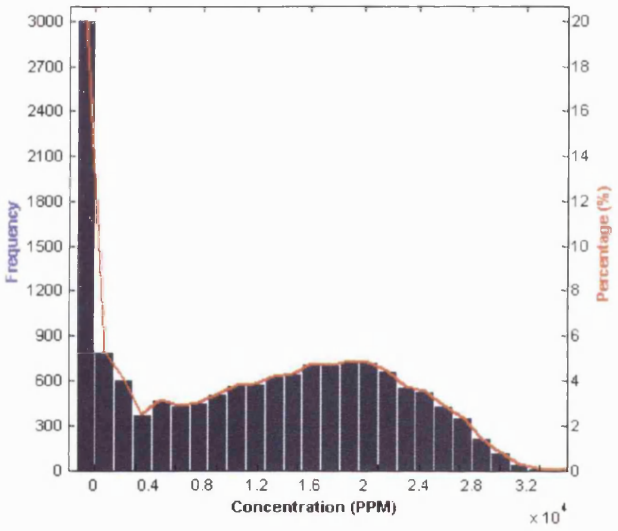


Figure 6.8: Frequency and percentage distribution of O_2

Note:
1. Number of blanks = 3004
2. Percentage data within the illustrated region = 99.86%
3. Maximum concentration of O_2 = 211118 PPM

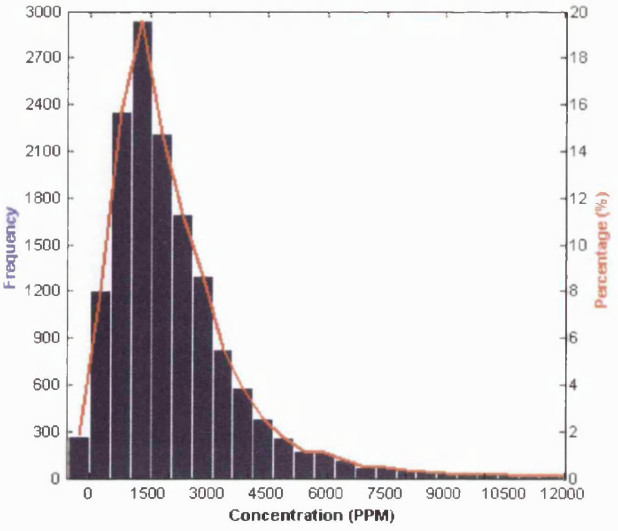


Figure 6.9: Frequency and percentage distribution of CO_2

Note:
1. Number of blanks = 267
2. Percentage data within the illustrated region = 99.47%
3. Maximum concentration of CO_2 = 78000 PPM

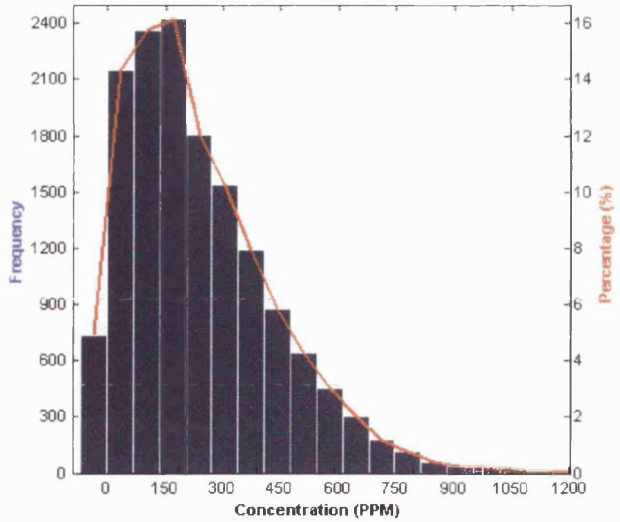


Figure 6.10: Frequency and percentage distribution of CO

Note:
1. Number of blanks = 735
2. Percentage data within the illustrated region = 99.79%
3. Maximum concentration of CO = 10255 PPM

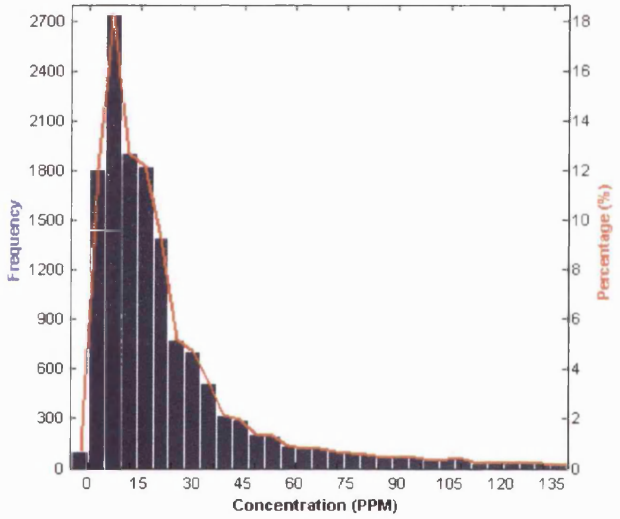


Figure 6.11: Frequency and percentage distribution of H_2

Note:
1. Number of blanks = 103
2. Percentage data within the illustrated region = 93.82%
3. Maximum concentration of H_2 = 2319 PPM

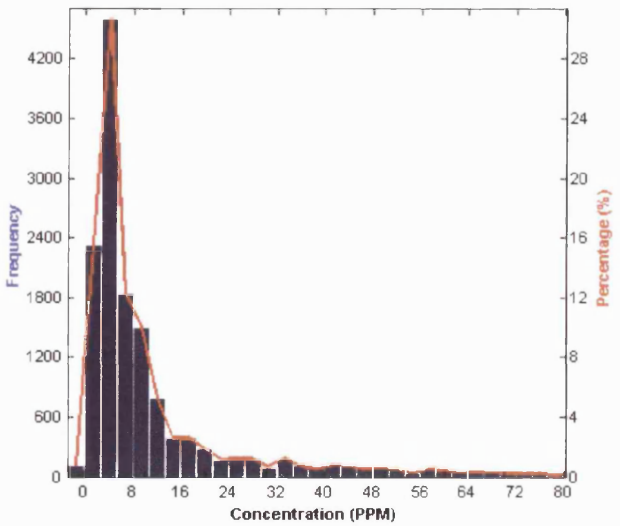


Figure 6.12: Frequency and percentage distribution of CH_4

Note:
1. Number of blanks = 103
2. Percentage data within the illustrated region = 94.29%
3. Maximum concentration of CH_4 = 2400 PPM

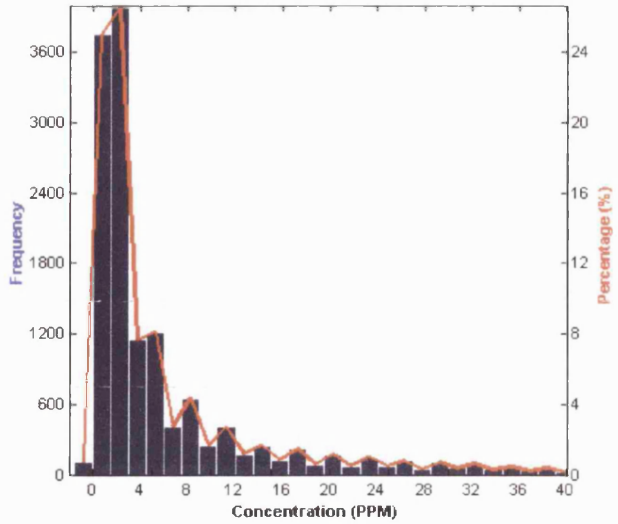


Figure 6.13: Frequency and percentage distribution of C_2H_6

Note:
1. Number of blanks = 103
2. Percentage data within the illustrated region = 94.29%
3. Maximum concentration of C_2H_6 = 1354 PPM

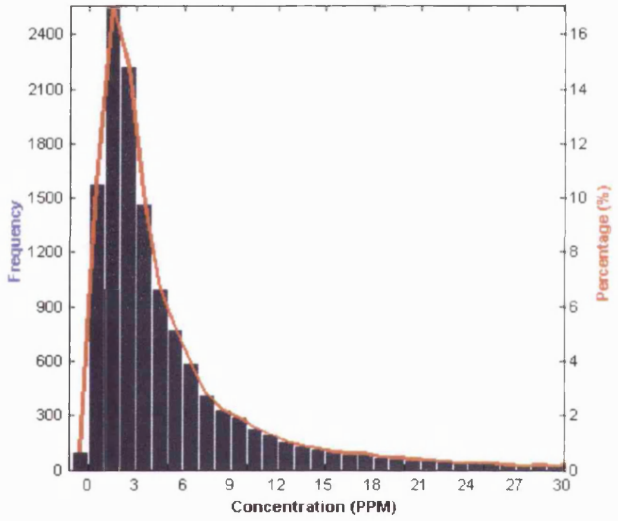


Figure 6.14: Frequency and percentage distribution of C_2H_4

Note:
1. Number of blanks = 103
2. Percentage data within the illustrated region = 86.78%
3. Maximum concentration of C_2H_4 = 2800 PPM

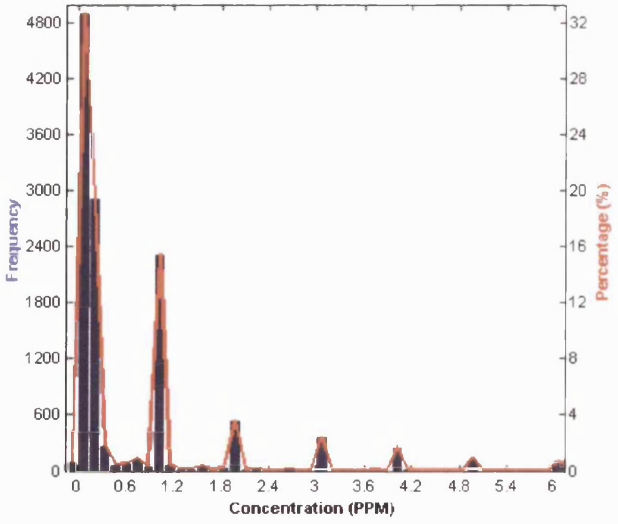


Figure 6.15: Frequency and percentage distribution of C_2H_2

Note:
1. Number of blanks = 104
2. Percentage data within the illustrated region = 83.67%
3. Maximum concentration of C_2H_2 = 1362.6 PPM

6.6 Percentage Distribution of Dissolved Gases According to Various Voltage Levels and Power Ratings

As described in Section 6.4, the acquired DGA database was divided into several subsets according to various voltage levels and power ratings. It is the objective of this section to conduct a meaningful comparison of the percentage distribution of various dissolved gases in these subsets of DGA data.

Since all subsets are of different sizes, it needs to be decided as to how comparison can be performed effectively. Some of the subsets, which are categorised under the voltage levels of 275/66 kV and 400/66 kV, and power ratings of 18 MVA, 210 MVA and 460 MVA, are simply too small in number for their percentage distributions to be displayed and compared effectively. Hence, the foregoing subsets were omitted from the comparison.

Some examples of the percentage distribution of dissolved gases according to various voltage levels and power ratings are illustrated in Figures 6.16 to 6.23. As can be observed from these figures, it can be concluded that the characteristic of the distribution is generally similar for the dissolved gases in power transformers of various voltage levels and power ratings, i.e. the dissolved gases are found to be normally concentrated in areas of low concentrations when compared with their maximum concentrations. Note that some discrepancies in distribution patterns may be observed due to the differences in the size of data and the designated range of the histogram.

In conclusion, the foregoing observations as reported in Section 6.5 and Section 6.6 correspond very well to the real-life circumstances in which the majority of DGA data being collected is from power transformers which are operating in a normal or faultless manner. In addition, the few cases of which high concentrations are found correspond to the situation in which some fault incidences have occurred inside the power transformers.

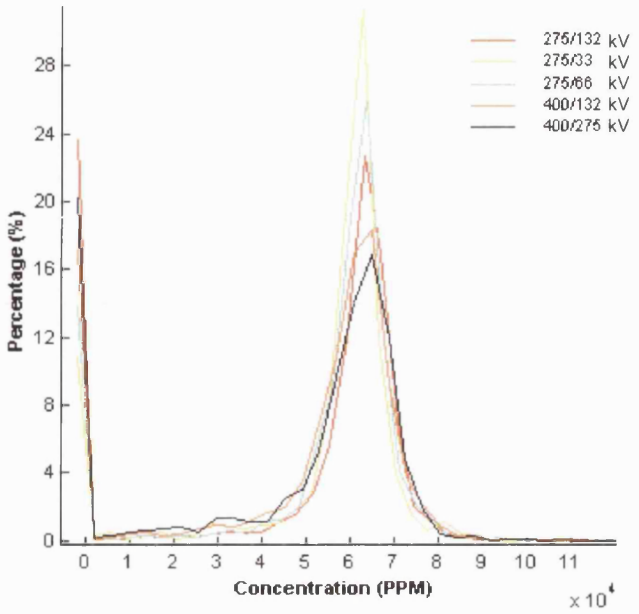


Figure 6.16: Percentage distribution of N_2 according to various voltage levels

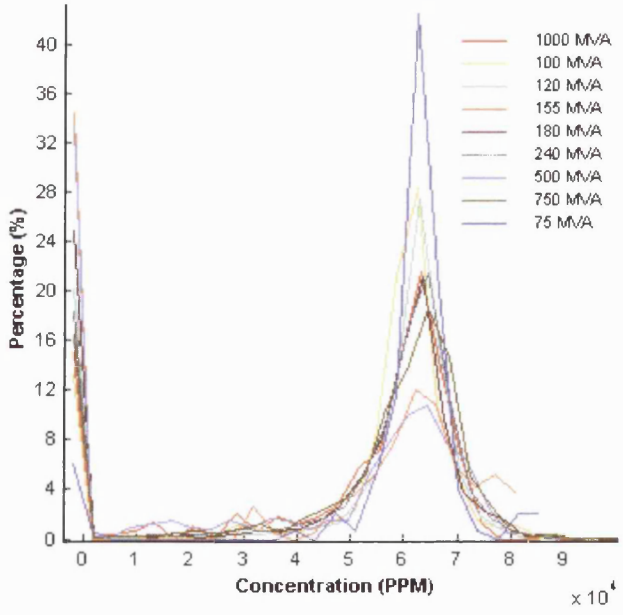


Figure 6.17: Percentage distribution of N_2 according to various power ratings

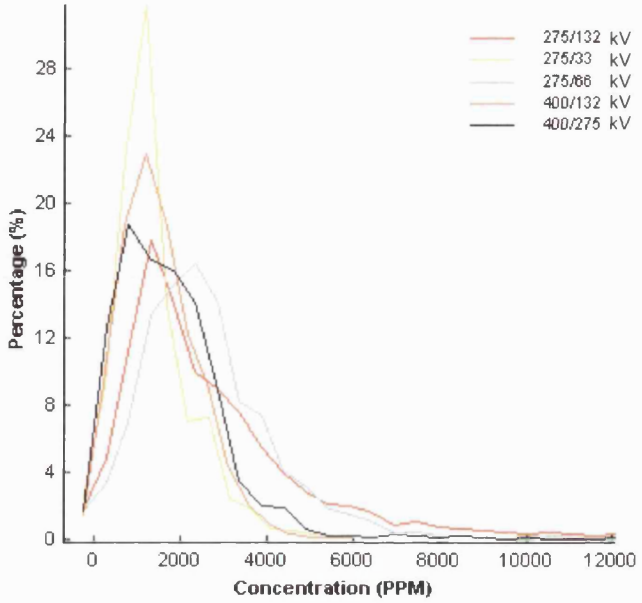


Figure 6.18: Percentage distribution of CO_2 according to various voltage levels

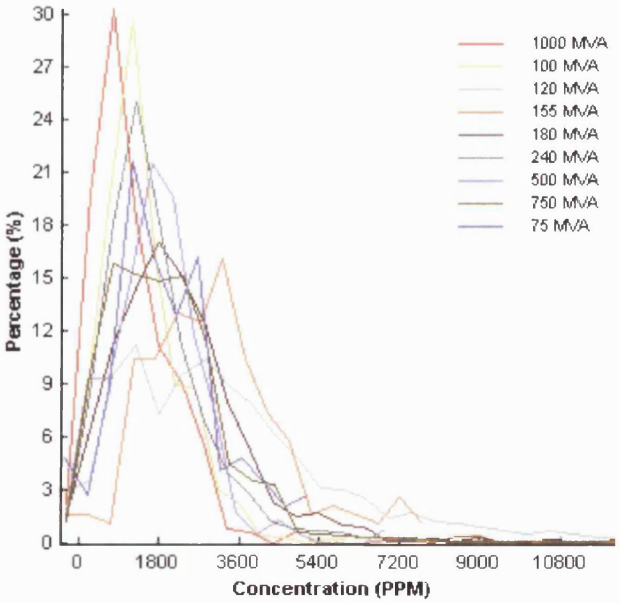


Figure 6.19: Percentage distribution of CO_2 according to various power ratings

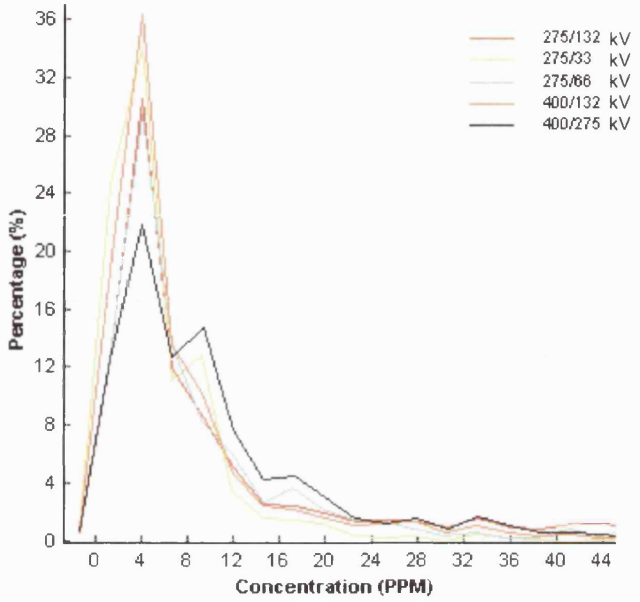


Figure 6.20: Percentage distribution of CH_4 according to various voltage ratios

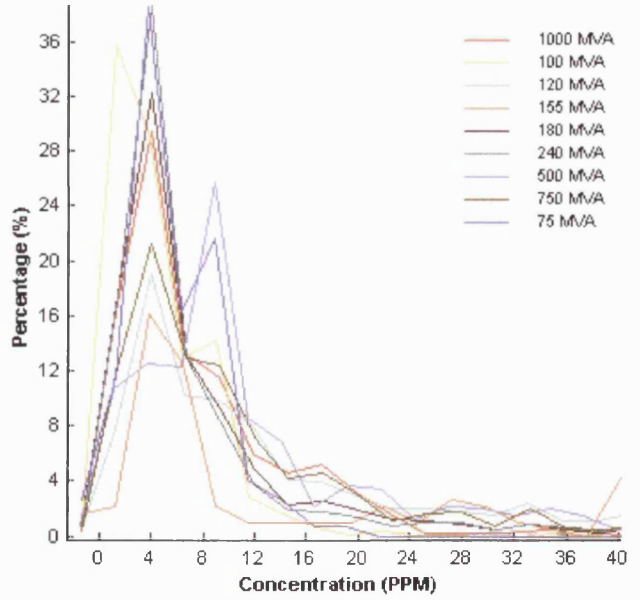


Figure 6.21: Percentage distribution of CH_4 according to various power ratings

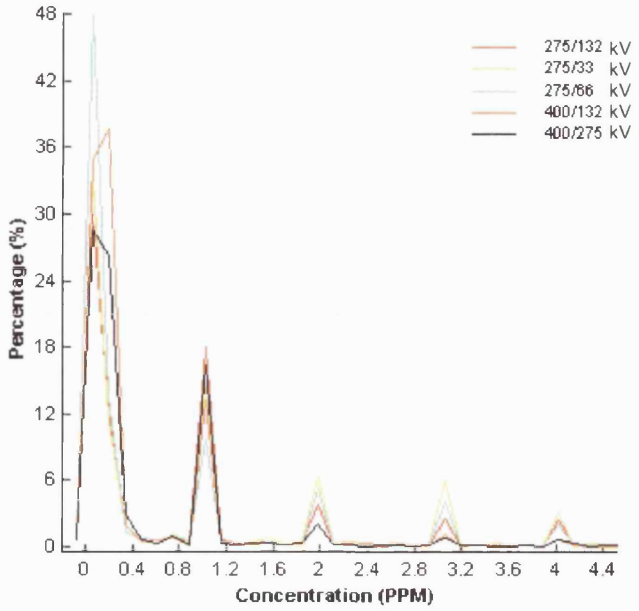


Figure 6.22: Percentage distribution of C_2H_2 according to various voltage levels

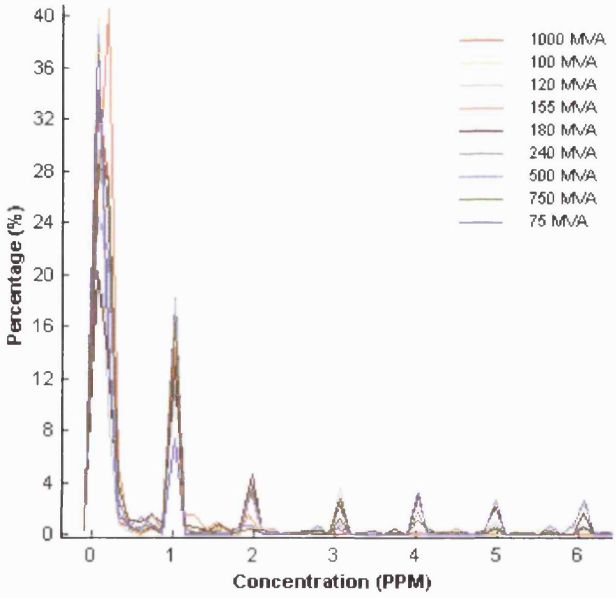


Figure 6.23: Percentage distribution of C_2H_2 according to various power ratings

6.7 Other Statistical Analyses Performed on DGA Data of Power Transformers

Apart from examining the partition of DGA records according to power ratings and voltage levels of power transformers, and the frequency and percentage distributions of each dissolved gas, other statistical measures and plots were also considered. Firstly, the measures of data range (i.e. the difference between maximum and minimum values), average and standard deviation were calculated for each dissolved gas. It was observed that sometimes large data range was obtained owing to the inclusion of abnormally high concentration values of certain dissolved gases; the presence of these "suspicious" DGA records might be due to human errors in data inputting. Moreover, the average and standard deviation may also be affected by the presence of these "suspicious" values. Therefore, appropriate measures must be taken to eliminate these values in subsequent higher-level of analysis on the DGA data.

Secondly, scatter diagrams were also plotted based on concentration of dissolved gases and their ratios. The reason for doing so is to determine the presence of clustering characteristics and to detect outliers within these scatter diagrams. It was subsequently observed that the majority of the data points would concentrate on certain region within the diagram, while other data points are located further away from the foregoing group of points and scattered around in either x- or y-axis. The former was thought to correspond to "normal" circumstances while the latter corresponds to more "rare" or unusual circumstances. Higher-level of analysis is necessary in order to establish the reasons for above observations.

6.8 Summary

Various statistical analyses have been conducted on the DGA data of power transformers. The DGA data was processed in order to remove several imperfections in the data. Besides, the processed DGA data was divided according to various

voltage levels and power ratings; analyses were conducted relating the amount of DGA data to the number and the total age of power transformers within each category of voltage levels and power ratings. Moreover, the frequency and percentage distribution of dissolved gases were also examined, which shows that the majority of dissolved gases are of low concentrations and there are only a few cases of which extremely high concentrations are found. The foregoing findings actually correspond well with the real-world circumstances in which power transformers are normally operating faultlessly and hence the recorded dissolved gases are usually of low concentrations. Hence, high concentrations of dissolved gases can usually be associated with abnormal or fault incidences that have occurred within the power transformers.

CHAPTER 7

DATA MINING ON THE DGA DATA OF POWER TRANSFORMERS USING THE SOM

7.1 Introduction

The proposed approach based on the self-organising map (SOM) was applied for the analysis and interpretation of the dissolved gas analysis (DGA) data of power transformers. Firstly, this chapter presents a feasibility study of the proposed approach, which was conducted on six selected subsets of the DGA data. Secondly, a further application of the proposed approach to the entire DGA database of power transformers is presented. Based on these investigations, an improved approach for interpretation of the DGA data is recommended, and it has been validated using conventional DGA interpretation schemes and actual fault cases.

7.2 The Proposed Approach

The proposed approach is fundamentally based on a data mining (DM) method known as the SOM, which has been introduced in Chapter 5. Basically, the proposed approach comprises of various stages, as illustrated in Figure 7.1. Note that there is a close resemblance between Figures 7.1 and 5.1 and this stems from the fact that the stage of DM using the SOM is an important constituent of the entire knowledge discovery in databases (KDD) process. However, all preceding and following stages of DM are also vital and careful consideration must thus be given to each of these stages so as to ensure the successful application of the proposed approach.

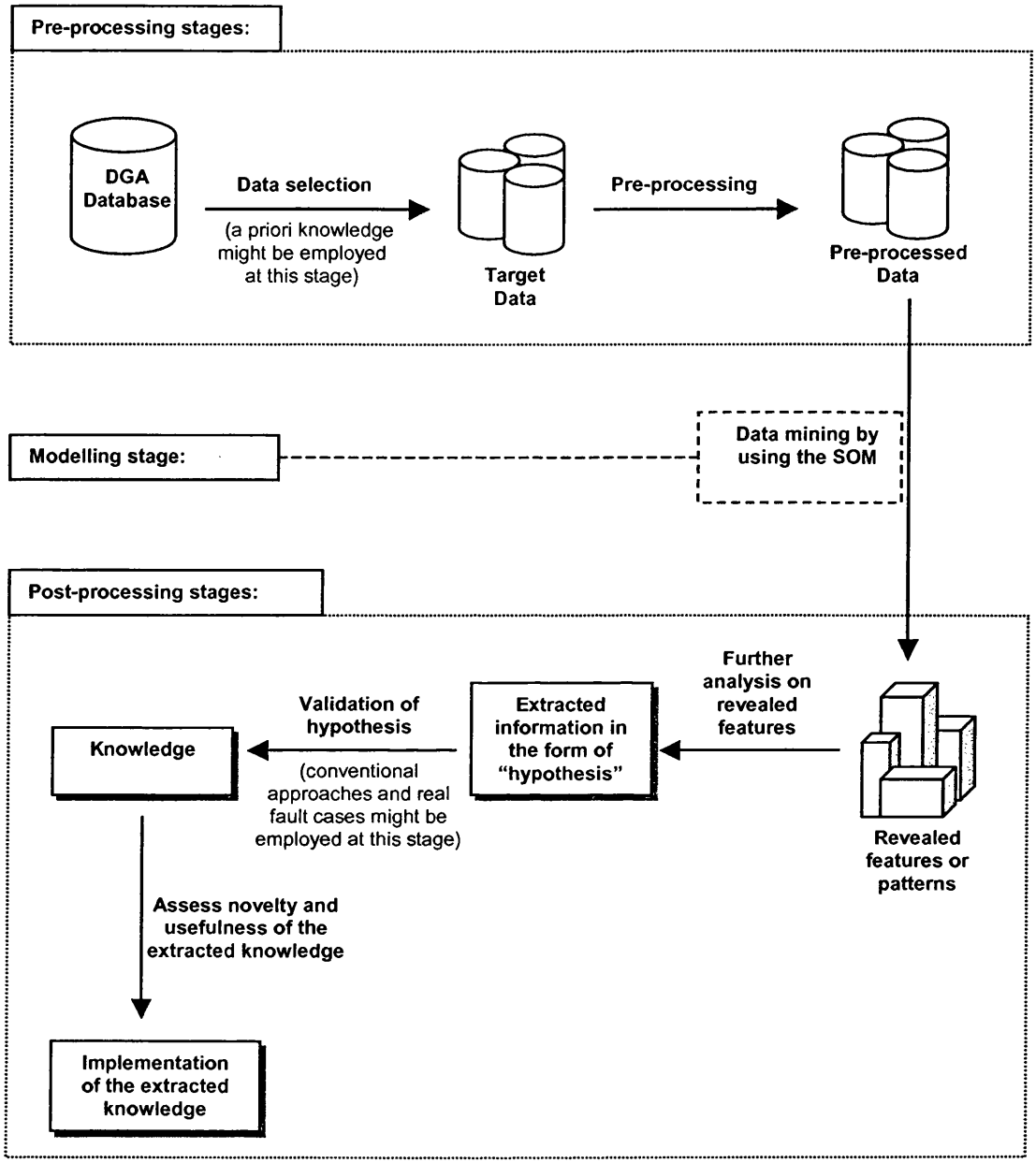


Figure 7.1: Various stages of the proposed approach

7.3 Justification for Collective Analysis of the DGA Data

It is well established that various parameters concerning the type and functionality of transformers actually influence the analysis and interpretation of the DGA data. Since the DGA database as obtained from the National Grid Company (NGC), UK,

corresponds to only substation power transformers in the NGC's transmission system, which according to [21] all belong to the same type of open system, core-type construction and with separate on-load tap changer (OLTC), the DGA data of the acquired database can thus be analysed and interpreted collectively.

It shall be mentioned that the foregoing means of analysing and interpreting the DGA data, i.e. according to the type and functionality of transformers, is also identical to the approach adopted by various international bodies such as the IEC [20]. Apart from the IEC Ratios, all other well-known conventional DGA schemes such as Dörnenburg Ratios [15], Rogers Ratios [16], Duval Triangle [17, 18] and CIGRÉ Methods [21] are also recommended to be implemented on power transformers without reference to certain voltage levels and power ratings.

In addition, it has been shown that the characteristic of statistical distribution of dissolved gases for the acquired DGA database is generally similar for various subsets of the DGA data, as explained in Section 6.6. Thus, it is justifiable to analyse the DGA data of power transformers of various voltage levels and power ratings in a collective manner, on condition that all power transformers are from the same type and performing an identical function.

Finally, since the effect of the volume of oil towards the volume of dissolved gases has been eliminated through the adoption of "part-per-million" (PPM) measurement of the concentration level (i.e. 1 PPM means 1 micro-litre of a dissolved gas in 1 litre of an oil sample), it is acceptable to analyse the DGA data from various power transformers in a collective manner regardless of the volume of transformer tanks; this is based on the fact that a fix amount of oil sample is always taken from the tank and assuming that dissolved gases have already been distributed evenly in the insulation oil when the sample was extracted from the bottom of the tank.

7.4 A Feasibility Study of the Proposed Approach

The feasibility study is based on six selected subsets of a total of 755 DGA records, which correspond to power transformers from three different manufacturers, of two different voltage levels and of one power rating, as illustrated in Table 7.1. Herein, these subsets of DGA data are collectively referred to as the "Sixsets" DGA data.

Table 7.1: The "Sixsets" DGA data

Subset	Manufacturer *	Voltage Level (kV)	Power Rating (MVA)	Number of Data
1	I	400/132	240	72
2	I	275/132	240	91
3	II	275/132	240	187
4	II	400/132	240	94
5	III	400/132	240	108
6	III	275/132	240	203

* Roman numerals are used to represent different manufacturers.

7.4.1 Pre-Processing of the "Sixsets" DGA Data

The "Sixsets" DGA data has to be pre-processed before being submitted for analysis using the SOM. Firstly, each "blank" (i.e. no data entry or record) in the "Sixsets" DGA data was replaced by the notation "NaN" (i.e. Not-a-Number) for reasons explained in Section 6.3. Likewise, each "0 PPM" record of N_2, O_2, CO_2 and CO was also replaced by the "NaN" again for reasons explained in Section 6.3. The SOM algorithm is capable of dealing with the "NaN" by ignoring those components of reference vectors which contain "NaNs" during the calculation of Euclidean distances throughout the training process. This is accomplished without affecting the ability of SOM to learn the inherent characteristics of the DGA data. Finally, those "$< x$" entries in the "Sixsets" DGA data were converted systematically to reasonably small values according to Table 6.2.

7.4.2 Data Mining on the "Sixsets" DGA Data

The pre-processed DGA data was submitted for analysis using the SOM. As mentioned in Section 6.2, each DGA data contains information on concentration of moisture (H_2O) and nine dissolved gases, i.e. N_2, O_2, CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2. Therefore, four sets of DGA data were constructed from the foregoing components and three scaling or transformation methods were applied to these data sets; sixteen distinct configurations of training data are thus resulted as a consequence of the foregoing, as illustrated in Table 7.2.

The basis for adopting various sensible combinations of input components and scaling methods is to search for interesting features or patterns in the optimum SOM for each configuration of the training data. Moreover, it is also the aim to determine those key or important dissolved gases that have actually contributed to the observed features or patterns in these optimum SOMs. Brief description on the "range", "variance" and "logarithm" scaling methods are given in the Appendix of this thesis.

Table 7.2: The "Sixsets" training data of SOM

Set	Input Components	Configuration of Training Data			
A [a]	H_2O, N_2, O_2, CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**1**: Not scaled	**2**: Range scaling	**3**: Variance scaling	**4**: Logarithm scaling
B [b]	CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**5**: Not scaled	**6**: Range scaling	**7**: Variance scaling	**8**: Logarithm scaling
C [c]	CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**9**: Not scaled	**10**: Range scaling	**11**: Variance scaling	**12**: Logarithm scaling
D [d]	H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**13**: Not scaled	**14**: Range scaling	**15**: Variance scaling	**16**: Logarithm scaling

a. All dissolved gases including moisture.
b. Key dissolved gases from the degradation of cellulose and insulation oil.
c. Dissolved gases that are combustible by nature.
d. Key dissolved gases from the degradation of insulation oil.

In addition, suitable configuration and parameters for the training of SOM were selected according to the guidance for practical application of SOM outlined in Section 5.5.2, as summarised in Table 7.3. Thus, a total of sixteen SOMs were trained based on the "Sixsets" DGA data that was configured according to Table 7.2. Finally, the optimum SOM for each configuration of training data was selected according to the procedures outlined in Section 5.5.2.4.

7.4.3 Chosen Optimum SOMs of the "Sixsets" DGA Data

Optimum SOMs of the "Sixsets" DGA data can be visualised using u-matrix and component-plane illustrations. Through observation of the features or patterns as displayed by these optimum SOMs, it was discovered that the application of different scaling methods has lead to huge differences in the visualisation quality of the trained SOMs.

Firstly, "masking" or "shadowing" of higher-valued dissolved gases, e.g. N_2 and O_2, over those lower-valued dissolved gases, e.g. C_2H_6 and C_2H_2, was apparent if no scaling was applied to the data set. Secondly, the use of "variance" scaling has resulted in "messy" maps, in which the inherent data characteristics were not very perceptible. Thirdly, the optimum SOMs for those data that were scaled using "logarithm" method were found to be very "messy" and ambiguous, where it is not possible to establish a clear picture on the inherent relationship among dissolved gases and their hidden structure. On the contrary, interesting and comprehensible features can be discerned for those SOMs that were trained using the DGA data that had been scaled by the "range" method, i.e. Configurations 2, 6, 10 and 14 of the training data. Thus, these optimum SOMs were chosen for further investigation.

Table 7.3: Configuration, training and visualisation parameters of SOM

No.		Phase 1 Rough Ordering	Phase 2 Fine Tuning
	Configuration of SOM		
1	Shape	Rectangular	
2	Arrangement of neurons	Hexagonal configuration	
3	Number of neurons	$20 \times \sqrt{dlen}$ [a]	
4	[*x-dim*, *y-dim*] [b]	$\frac{x-dim}{y-dim} = \sqrt{\frac{\alpha}{\beta}}$ [c]	
	Training Parameters		
5	Neighbourhood function	Rectangular	
6	Training method	Sequential	
7	Radius	Max([*x-dim*, *y-dim*])	1/8 × Radius_Phase 1
8	Learning rate	0.05	0.001
9	Learning rate function	Linear	
10	Training iterations	According to Section 5.5.2.4	
	Visualisation Parameter		
11	Function for u-matrix calculation	Mean (see Section 5.5.2.3)	

a. *dlen*: Total number of input vectors.
b. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
c. α, β: First and second eigenvalues of the covariance matrix of input data.

In essence, the "range" method effectively transforms the minimum and maximum values of every input component into 0 and 1 respectively; all values in between are scaled according to Equation (1) as illustrated in the Appendix. The reason for using the "range" method is to limit the learning ranges of the SOM owing to the huge differences in concentration of various dissolved gases. Therefore, it is not surprising that excellent visualisation of features or patterns are observed since the concentration of all gases will only vary from 0 to 1, thereby imposing equal influences during the training process of SOM. Optimum SOMs for Configurations 2, 6, 10 and 14 of the training data are illustrated in Figures 7.2 to 7.9. Note that a 64-level colour scale was utilised for the illustration of component planes while a 64-level grey scale was utilised for the illustration of u-matrices for ease of visualisation. The map sizes, number of training iterations and SOM quality measures for each optimum SOM are presented in Table 7.4.

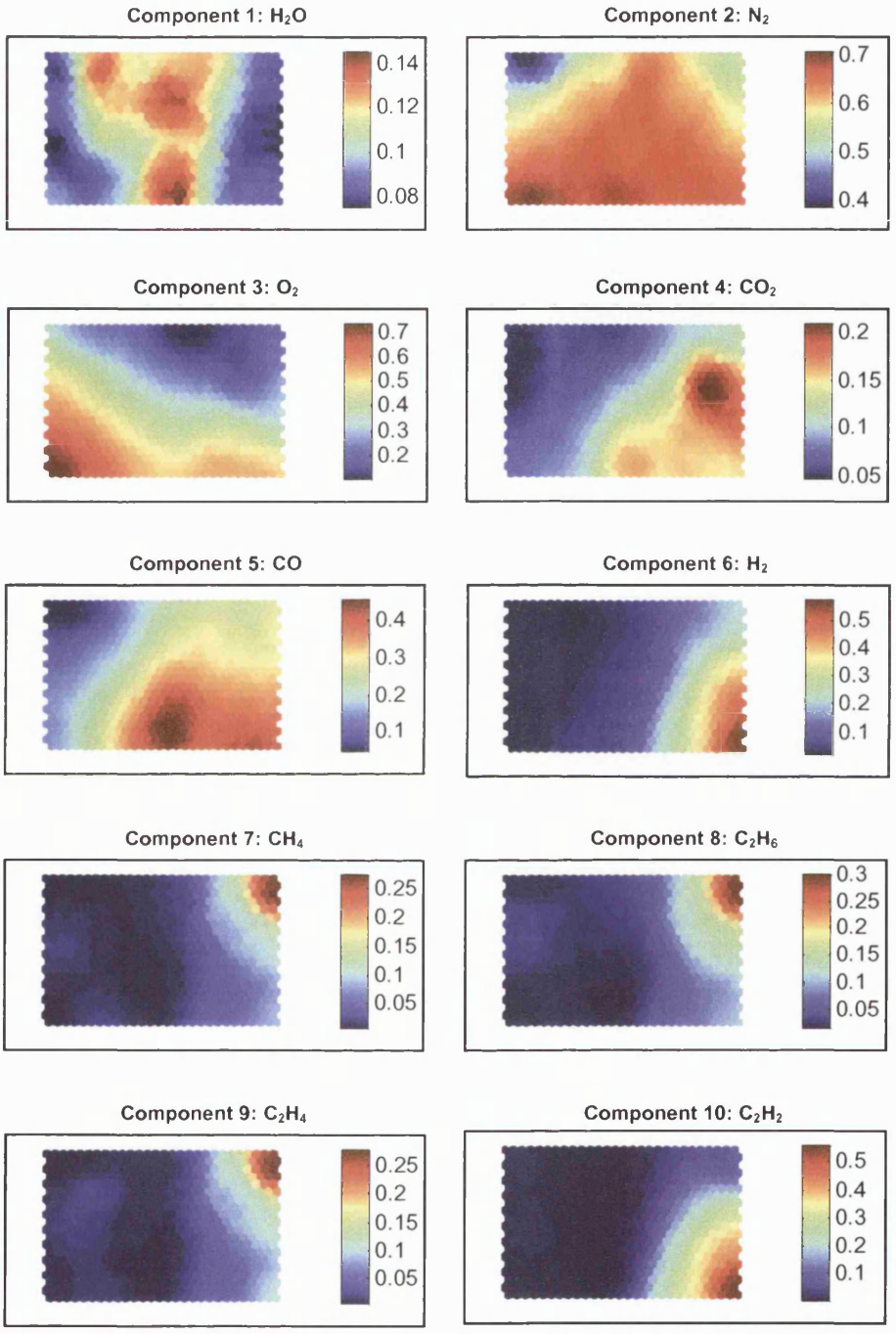


Figure 7.2: Component planes of the optimum SOM for Configuration 2 of the training data (i.e. Set A with "range" scaling)

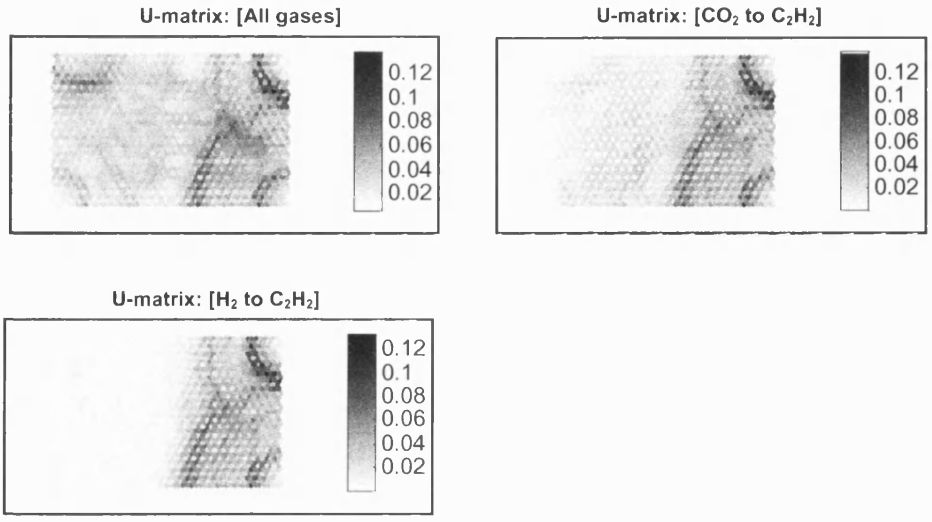


Figure 7.3: U-matrices of the optimum SOM for Configuration 2 of the training data (i.e. Set A with "range" scaling)

Table 7.4: The map sizes, number of training iterations and SOM quality measures for optimum SOMs of the "Sixsets" DGA data

No.	Training Data	Map Size [a] [*x-dim*, *y-dim*]	Number of Training Iterations	SOM Quality Measures [d]	
				AQE [b]	TE [c]
1	Configuration 2	[28 20]	22650 (30 passes of all training vectors)	0.1647	0.0570
2	Configuration 6	[34 16]	30200 (40 passes of all training vectors)	0.0885	0.0291
3	Configuration 10	[34 16]	22650 (30 passes of all training vectors)	0.0734	0.0636
4	Configuration 14	[30 18]	30200 (40 passes of all training vectors)	0.0454	0.0623

a. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
b. AQE: Average quantisation error
c. TE: Topographic error
d. AQE and TE at the specified number of training iterations during the fine-tuning phase.

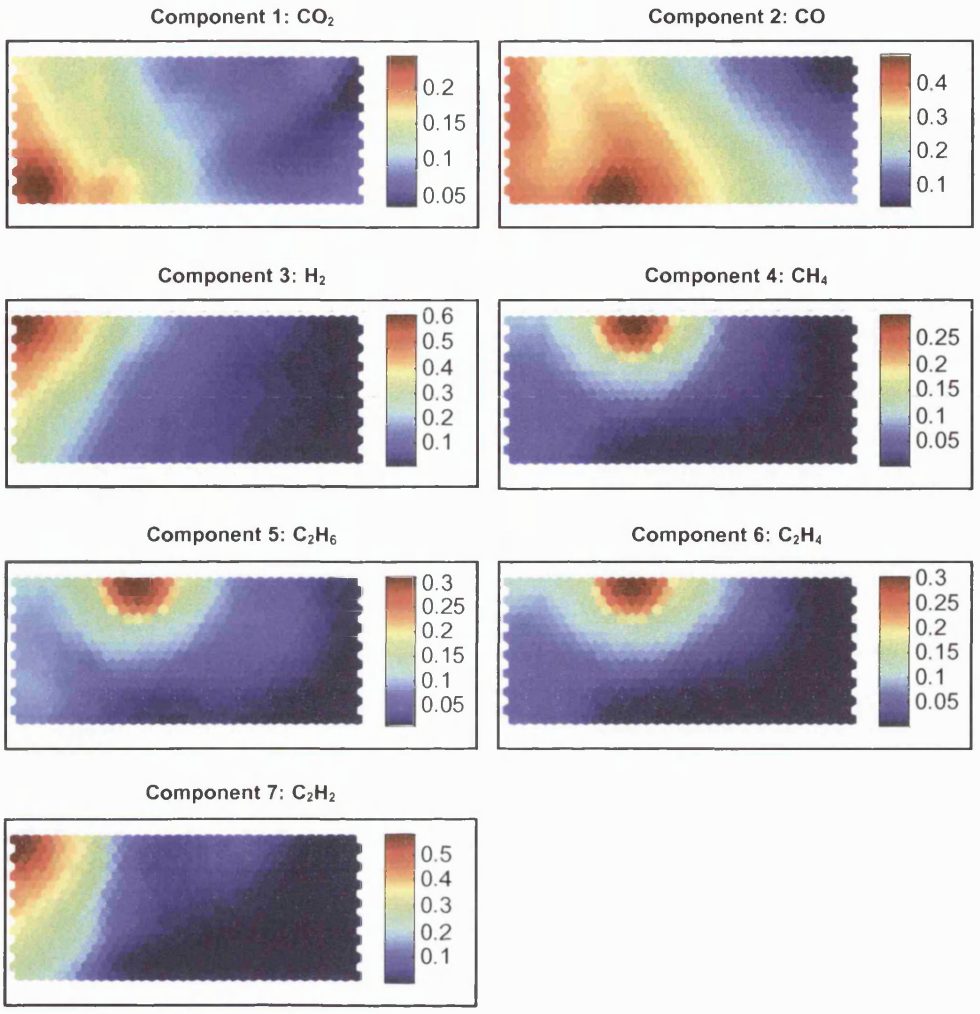


Figure 7.4: Component planes of the optimum SOM for Configuration 6 of the training data (i.e. Set B with "range" scaling)

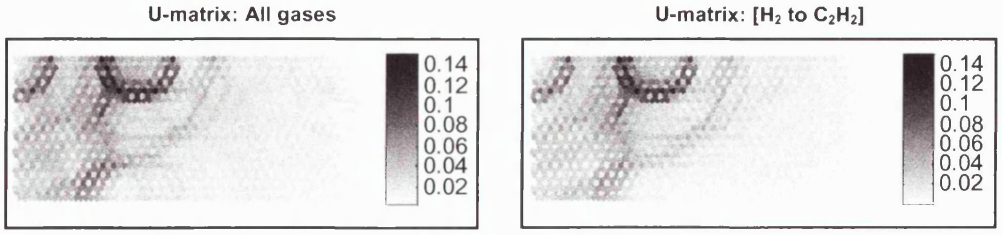


Figure 7.5: U-matrices of the optimum SOM for Configuration 6 of the training data (i.e. Set B with "range" scaling)

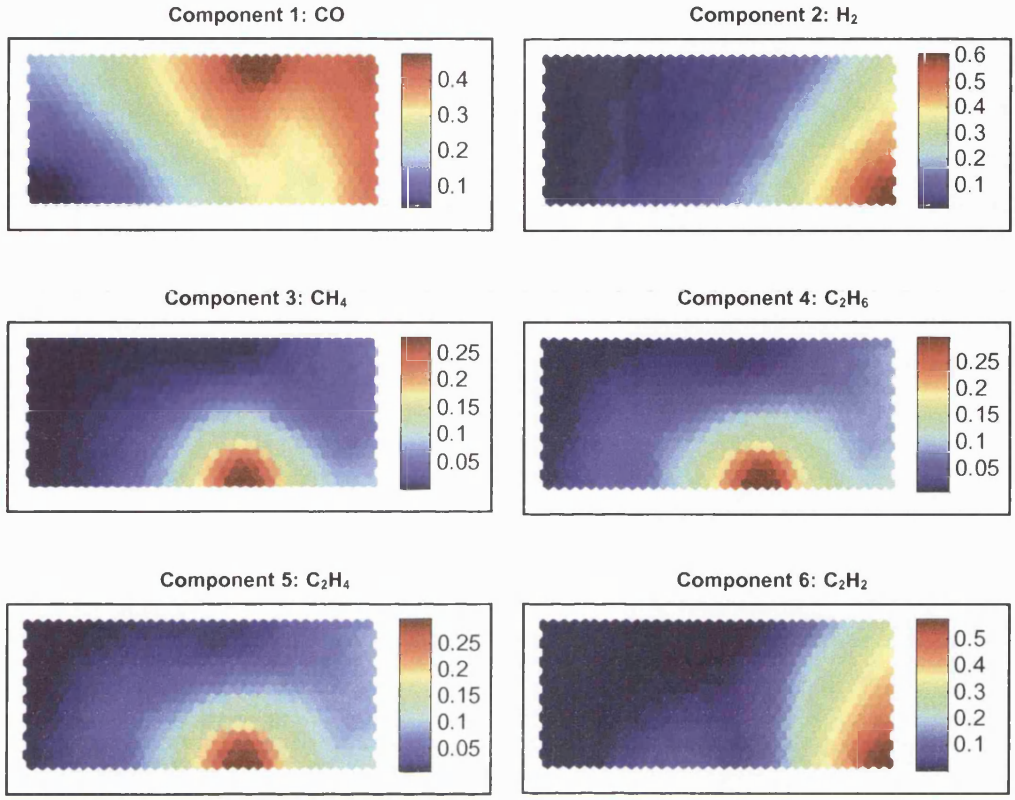


Figure 7.6: Component planes of the optimum SOM for Configuration 10 of the training data (i.e. Set C with "range" scaling)

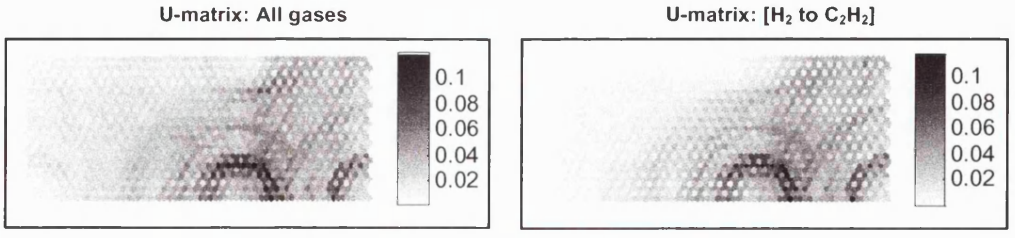


Figure 7.7: U-matrices of the optimum SOM for Configuration 10 of the training data (i.e. Set C with "range" scaling)

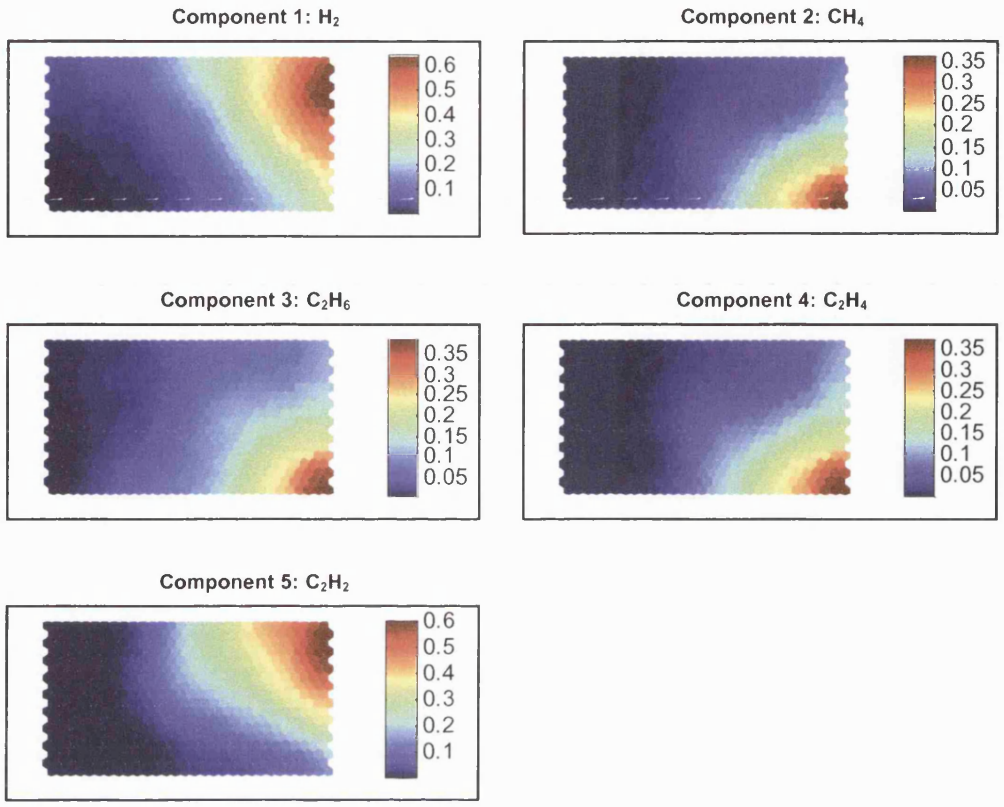


Figure 7.8: Component planes of the optimum SOM for Configuration 14 of the training data (i.e. Set D with "range" scaling)

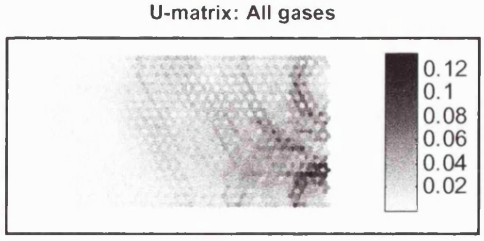


Figure 7.9: U-matrix of the optimum SOM for Configuration 14 of the training data (i.e. Set D with "range" scaling)

As explained in Section 5.5.1, every neuron of SOM is represented by an *n*-dimensional reference vector, where *n* is the dimension of the input space (i.e. number of components within an input vector). For this reason, there is a one-to-one correspondence between components of reference vectors and input components of training vectors. For example, the first component of all reference vectors of the optimum SOM for Set A actually corresponds to H_2O, the second component corresponds to N_2 and so on. Therefore, if reference vectors of the optimum SOM are "sliced" and viewed component-by-component, as shown by the illustration of component planes, then the inherent characteristics of input components can be effectively visualised.

Component planes for various chosen optimum SOMs are illustrated in Figures 7.2, 7.4, 7.6 and 7.8. Notice that the colourbar on the right-hand-side of every component plane illustrates the magnitude of scaled values that are learned by the SOM; the actual magnitude of values, which is measured in PPM, can be obtained by reversing the "range" scaling process. Although scaled values of the training data vary from 0 to 1, the SOM is found to be capable of learning about 90% of the training data in general, due to the fact that there is only a few training cases with extremely high concentration of dissolved gases. Therefore, the maximum magnitude, as shown in the colourbar, typically will not reach the value of 1. Nevertheless, the SOM has impeccably performed its function of "mining" or unearthing the inherent characteristics of the DGA data, which are intrinsically represented by the majority (over 90% in this case) of the training data.

As observed from Figures 7.2, 7.4, 7.6 and 7.8, close correlation is observed for two categories of dissolved gases, and this is signified by the similarity in the shape and location of revealed patterns on the component planes. The first category of correlated dissolved gases comprises of H_2 and C_2H_2 while the second category comprises of CH_4, C_2H_6 and C_2H_4. Besides, no apparent correlation is observed for all other dissolved gases. Nonetheless, it is observed that [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4] patterns do correspond to significant concentrations of CO_2 and CO.

Furthermore, the "spread" or distribution of dissolved gases with various levels of concentration can also be discerned from the component planes. It can be observed from the latter that the moisture, atmospheric gases (i.e. N_2, O_2 and CO_2), and products of cellulose degradation (i.e. CO_2 and CO) have a larger "spread" or distribution of concentration when compared with those "fault" gases, i.e. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2.

In actual fact, the foregoing observations are comprehensible since the concentration of "fault" gases is usually very low in power transformers that are operating faultlessly during most of their lifetime. Meanwhile, it is not uncommon to find higher and more distributed concentration in atmospheric gases and products of cellulose degradation for the reasons explained in Section 6.3. Therefore, it can be assumed that the half- and quarter-circular patterns of "fault" gases as discerned from the component planes actually resemble those circumstances where some problems or incipient faults are likely to have occurred within the power transformers, thereby giving rise to the high concentration of these gases.

In addition, the u-matrix illustration can be utilised for summarising the inherent characteristics of the DGA data, as shown in Figures 7.3, 7.5, 7.7 and 7.9. Notice that the Euclidean distances between SOM neurons are indicated by the grey colorbar on the right-hand-side of each u-matrix. As can be seen from these figures, the u-matrix can be calculated based on all components or selected components of the reference vectors. Nevertheless, at all instances, two groups of apparent co-centre clusters are always observed on every u-matrix illustration, which actually correspond to the two categories of correlated dissolved gases, i.e. [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4]. In addition, these two groups of clusters also intersect with one another, forming another interesting cluster on the u-matrix illustrations.

Judging from the foregoing observations on component planes and u-matrices of chosen optimum SOMs for the "Sixsets" DGA data, it can be concluded that the "key" dissolved gases of power transformers, which have contributed to the observed features or patterns, are the five "fault" gases, i.e. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2.

7.4.4 Analysis on Revealed Features of the "Sixsets" DGA Data

As reported in Section 7.4.3, the SOM has successfully learned the inherent characteristics of the "Sixsets" DGA data and presented them in a comprehensible and discernible format. Hitherto, it is known that both H_2 and C_2H_2 are correlated to each other; the same also applies to CH_4, C_2H_6 and C_2H_4. But, what do the foregoing observations actually mean in the context of real-world circumstances?

Generally, a large quantity of C_2H_2 will be generated and dissolved in the insulation oil during the onset and evolution of discharges such as arcing, sparking etc. In addition, H_2 will also be generated in abundance as a direct consequence of the nature of high heat and high temperature of this incipient fault. Therefore, the observed correlation between H_2 and C_2H_2 is easily comprehensible. In the case of thermal fault (TF), such as overheating or hotspots, it is known that large quantities of CH_4, C_2H_6 and C_2H_4 will be generated, the amount of which depends on the intensity of the fault, and this is signified by the fault temperature. Thus, the observed correlation among CH_4, C_2H_6 and C_2H_4 is also understandable. Besides, if the aforementioned faults also involve the cellulose insulation, both CO_2 and CO will also be produced along side the foregoing "fault" gases. Nevertheless, although the concentration of CO_2 and CO is quite high in regions of high concentration of "fault" gases, no apparent correlation is revealed. The main emphasis of subsequent analyses will thus be based on the five "fault" gases.

Further statistical analysis was conducted on chosen optimum SOMs of the "Sixsets" DGA data so as to determine the following:

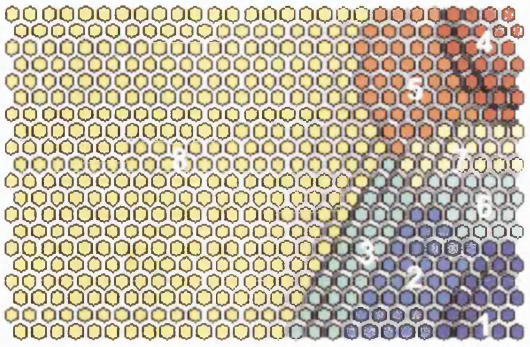
- Whether the observed correlation in two categories of dissolved gases can be proven through statistical means?
- Whether those clusters in each group of co-centre clusters, as observed on the u-matrices of the optimum SOMs, are actually similar in characteristics?

Essentially, the analysis is comprised of the following steps:

- Step 1: Identification of clusters through colour coding on u-matrices of chosen optimum SOMs.
- Step 2: Identification of those training vectors that regard the SOM neurons in each colour-coded cluster as their best matching units (BMUs).
- Step 3: Calculation of average concentration (measured in PPM) of key dissolved gases, i.e. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, for each colour-coded cluster from the training vectors gathered in Step 2.

Apart from the analysis on identified clusters, regions which are formed with no apparent clusters by the key dissolved gases were also examined and similar steps of analysis were performed on these regions so as to determine the average concentration of key dissolved gases in these regions. Results of the foregoing analyses are illustrated in Figures 7.10 to 7.13. Several important characteristics can be summarised from these figures:

- Two distinct distribution patterns of key dissolved gases can be identified, which correspond to those of [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4], respectively.
- Among key dissolved gases, H_2 and C_2H_2 are dominant dissolved gases in the co-centre clusters formed by these two gases while CH_4, C_2H_6 and C_2H_4 are dominant dissolved gases in the co-centre clusters formed by these three gases.
- Average concentration of dominant dissolved gases in the respective group of co-centre clusters is also found to increase from the outer cluster to inner cluster.
- Those clusters in each group of co-centre clusters are found to have similar patterns of the composition of key dissolved gases.
- Interception clusters are found to have lower average concentration of C_2H_2 but higher average concentration of CH_4, C_2H_6 and C2H4 when compared with the [H_2 C_2H_2] clusters of the same levels. For example in Figure 7.11, Region 7 contains lower average concentration of C_2H_2 and higher average concentration of CH_4, C_2H_6 and C2H4 if compared to Region 2.



Colour coding:

1. 1, 2, 3: [H_2 C_2H_2] clusters
2. 4, 5: [CH_4 C_2H_6 C_2H_4] clusters
3. 6, 7: Interception clusters
4. 8: Apparent cluster not identified

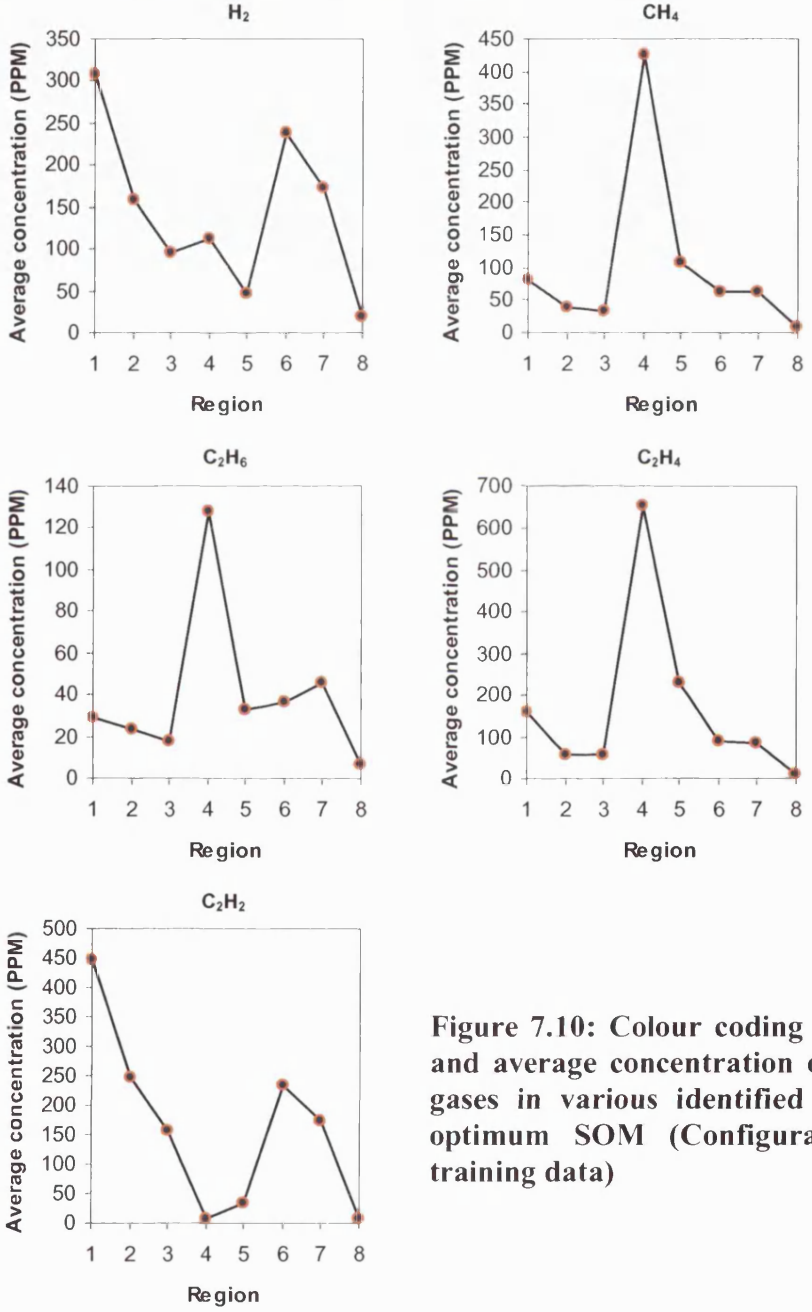


Figure 7.10: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 2 of the training data)

Colour coding:

1. 1, 2, 3, 4: [H_2 C_2H_2] clusters
2. 5, 6: [CH_4 C_2H_6 C_2H_4] clusters
3. 7, 8: Interception clusters
4. 9: Apparent cluster not identified

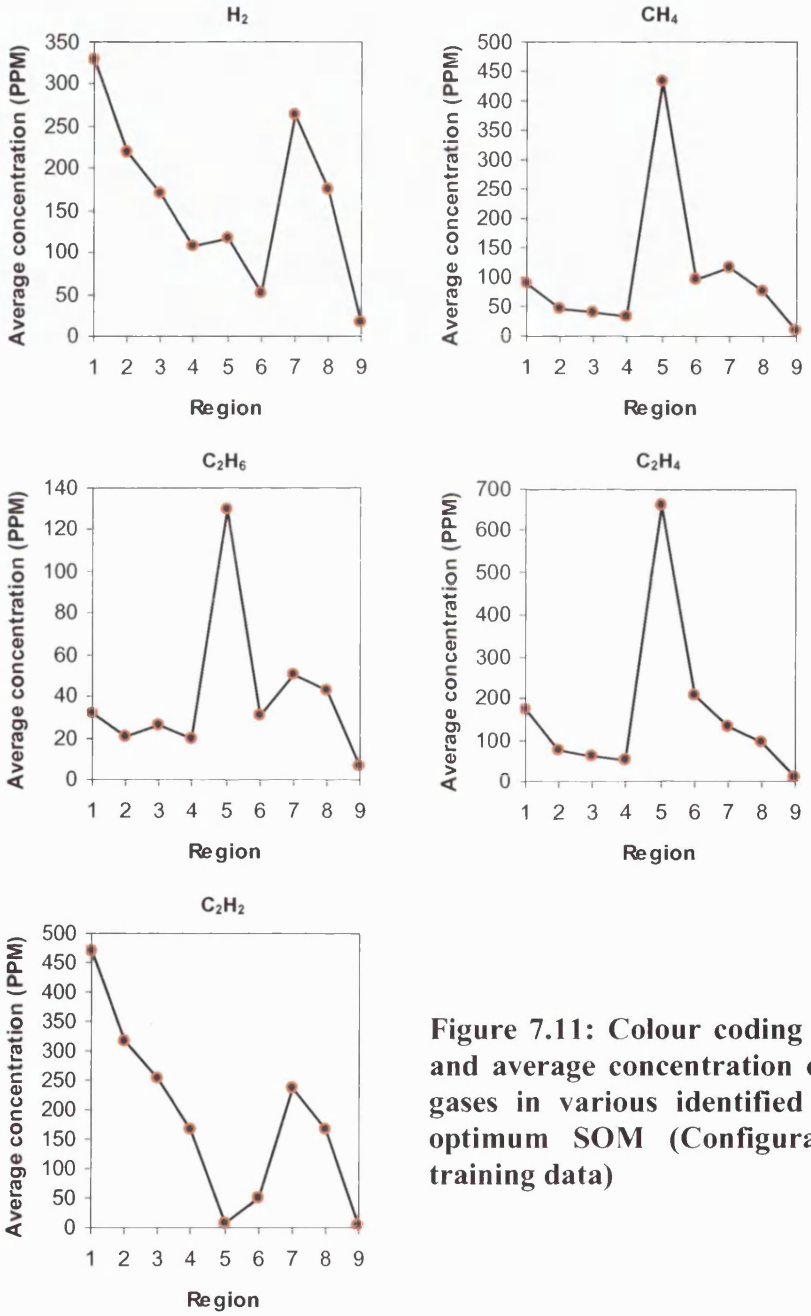
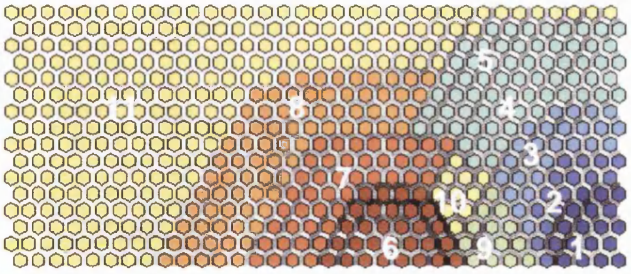


Figure 7.11: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 6 of the training data)



Colour coding:

1. 1, 2, 3, 4, 5: [H_2 C_2H_2] clusters
2. 6, 7, 8: [CH_4 C_2H_6 C_2H_4] clusters
3. 9, 10: Interception clusters
4. 11: Apparent cluster not identified

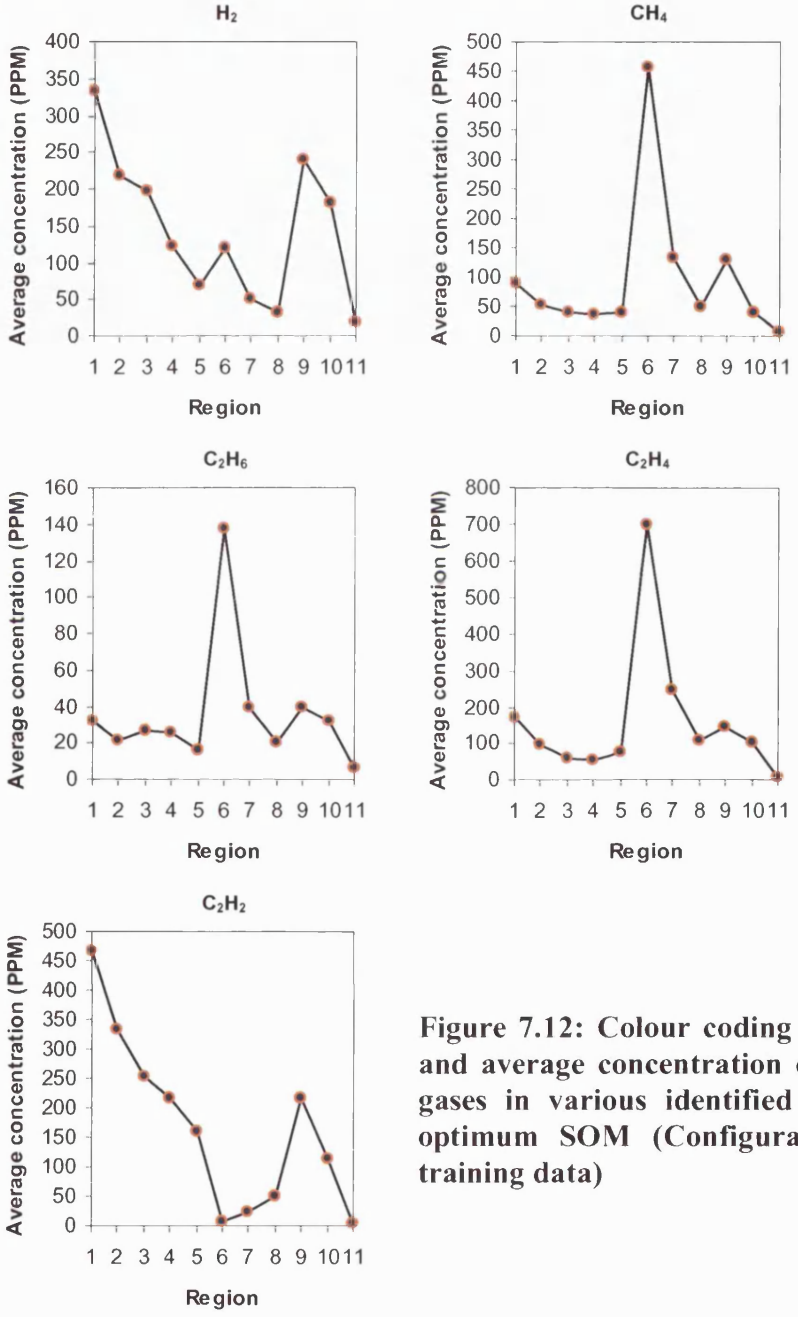
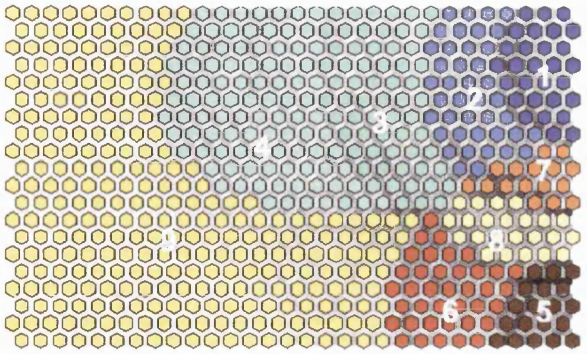


Figure 7.12: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 10 of the training data)



Colour coding:

1. 1, 2, 3, 4: [H_2 C_2H_2] clusters
2. 5, 6: [CH_4 C_2H_6 C_2H_4] clusters
3. 7, 8: Interception clusters
4. 9: Apparent cluster not identified

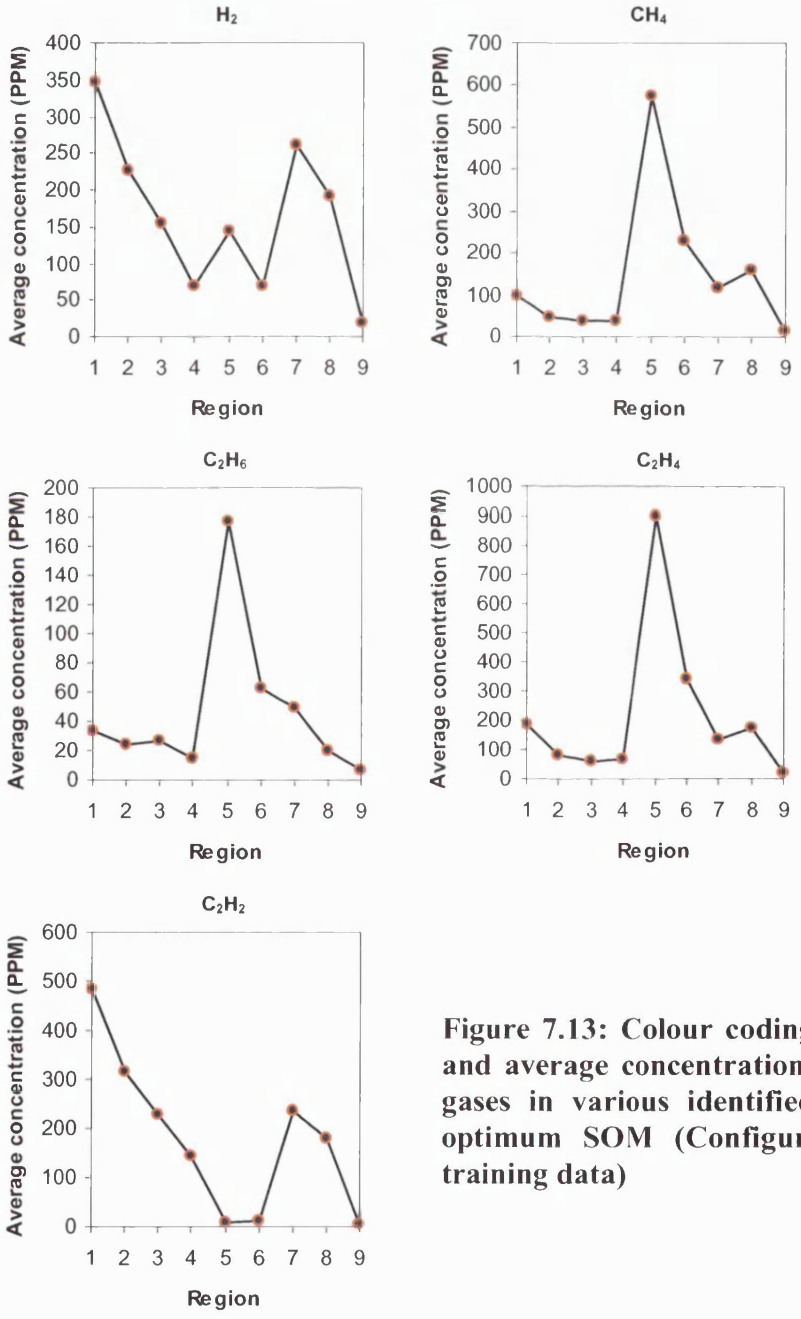


Figure 7.13: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 14 of the training data)

Several conclusions can thus be drawn from foregoing observations. Firstly, the observed correlation in [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4] has been validated owing to the similarity in the distribution pattern of average concentration across various clusters. Secondly, those clusters in each group of co-centre clusters are similar in characteristics, with gradual increase of the average concentration of dominant dissolved gases from the outer clusters to inner clusters. Thirdly, interception clusters are found to contain the characteristics of both [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4] clusters. Finally, those regions of which no apparent clusters are formed by key dissolved gases are assumed to be associated with the normal operating condition of power transformers, owing to the consistently low average concentration of key dissolved gases.

7.4.5 Hypothetical Association of Revealed Features

Owing to the a priori knowledge on the formation of dissolved gases during the onset and evolution of incipient faults and the outcome of the detailed statistical analysis on chosen optimum SOMs reported in Section 7.4.4, association can hence be hypothetically established between revealed features or patterns reported in Section 7.4.3 and several conditions of power transformers.

Two examples of the hypothesis, which are based on optimum SOMs of Configurations 6 and 10 of the training data, are illustrated in Figures 7.14 and 7.15. As can be observed from these figures, the "discharges" region corresponds to the group of co-centre clusters of which H_2 and C_2H_2 are found to be the dominant dissolved gases. In contrast, the "thermal fault" (TF) region corresponds to the group of co-centre clusters of which CH_4, C_2H_6 and C_2H_4 are found to be the dominant dissolved gases. These aforementioned groups of co-centre clusters also intercept with one another, forming another region that is assumed to be associated with the simultaneous occurrence of discharges and TF. Finally, the region in which no apparent clusters are formed by key dissolved gases is assumed to be associated with normal operating condition of power transformers, due to the consistently low average concentrations of all key dissolved gases.

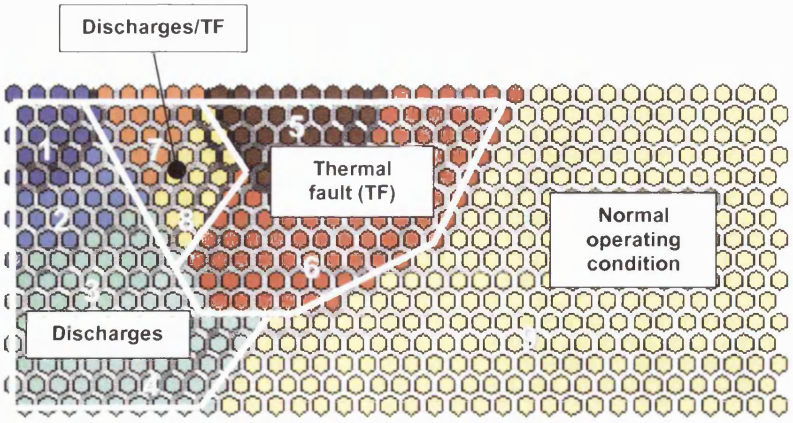


Figure 7.14: Association of revealed features with conditions of power transformers (based on Configuration 6 of the training data)

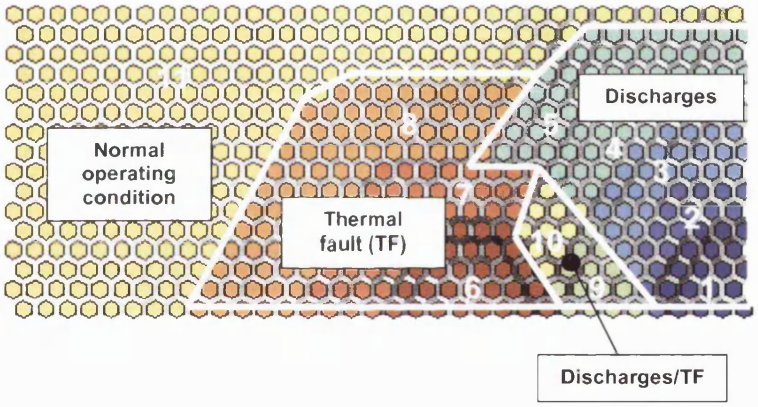


Figure 7.15: Association of revealed features with conditions of power transformers (based on Configuration 10 of the training data)

7.4.6 Validation of the Hypothesis

The hypothetical association of revealed features with conditions of power transformers, as presented in the previous section, has been validated using following approaches:

- Comparison with the interpretation as provided by several established conventional DGA interpretation schemes.
- Validation by using several actual fault cases.

7.4.6.1 Comparison with conventional DGA interpretation schemes

Training vectors that correspond to each numbered region illustrated in Figures 7.14 and 7.15 were gathered via the best-matching means and subsequently interpreted by conventional DGA schemes after being scaled back to original values. It shall be emphasised that these training vectors are actual DGA records of power transformers. Results of the foregoing are then compared with the interpretation as provided by the hypothesis illustrated in Figures 7.14 and 7.15. Since the type and severity of incipient faults detectable by conventional DGA interpretation schemes vary from one scheme to another, it is desirable to simply identify four main conditions of power transformers, i.e. normal operation, discharges, partial discharges (PD) and thermal fault (TF). Following conventional DGA interpretation schemes were utilised for the comparison:

- Dörnenburg Ratios
- Rogers Ratios
- IEC-1999 Ratios
- Duval Triangle (1993-edition)

Note that only the latest version of IEC Ratios and Duval Triangle are considered for comparison purposes. In addition, CIGRÉ methods are not considered due to the complexity of the interpretation associated with this scheme.

Results of the comparison are shown in Tables 7.5 to 7.8 for the optimum SOM of Configuration 6 of the training data (hypothesis illustrated in Figure 7.14), and in Tables 7.9 to 7.12 for the optimum SOM of Configuration 10 of the training data (hypothesis illustrated in Figure 7.15).

Table 7.5: Comparison of hypothesis with Dörnenburg Ratios (correspond to Figure 7.14)

		Interpretation by Dörnenburg Ratios					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I [a]**	**N/A [b]**
1 [53]	Discharges (D)		53				
2 [25]	Discharges (D)		18				7
3 [20]	Discharges (D)		11				9
4 [62]	Discharges (D)	4	4			6	48
5 [32]	Thermal fault (TF)				32		
6 [34]	Thermal fault (TF)				13	17	4
7 [4]	Discharges/TF (D/TF)		1				3
8 [2]	Discharges/TF (D/TF)		2				
9 [523]	Normal	442				1	80

a. N/I: No interpretation.
b. N/A: Not applicable.

Table 7.6: Comparison of hypothesis with Rogers Ratios (correspond to Figure 7.14)

		Interpretation by Rogers Ratios				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I ***
1 [53]	Discharges (D)		53			
2 [25]	Discharges (D)		25			
3 [20]	Discharges (D)		19			1
4 [62]	Discharges (D)	3	44		1	14
5 [32]	Thermal fault (TF)				6	26
6 [34]	Thermal fault (TF)		1		20	13
7 [4]	Discharges/TF (D/TF)		2			2
8 [2]	Discharges/TF (D/TF)		2			
9 [523]	Normal	481	10		20	12

* N/I: No interpretation.

Table 7.7: Comparison of hypothesis with IEC-1999 Ratios (correspond to Figure 7.14)

		Interpretation by IEC-1999 Ratios				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I ***
1 [53]	Discharges (D)		53			
2 [25]	Discharges (D)		25			
3 [20]	Discharges (D)		18			2
4 [62]	Discharges (D)	3	45		1	13
5 [32]	Thermal fault (TF)				32	
6 [34]	Thermal fault (TF)		2		16	16
7 [4]	Discharges/TF (D/TF)		4			
8 [2]	Discharges/TF (D/TF)		2			
9 [523]	Normal	481	8		26	8

* N/I: No interpretation.

Table 7.8: Comparison of hypothesis with Duval Triangle (1993-edition) (correspond to Figure 7.14)

		Interpretation by Duval Triangle (1993-edition)				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**D/TF**
1 [53]	Discharges (D)		53			
2 [25]	Discharges (D)		25			
3 [20]	Discharges (D)		20			
4 [62]	Discharges (D)	3	56		3	
5 [32]	Thermal fault (TF)				32	
6 [34]	Thermal fault (TF)		6		19	9
7 [4]	Discharges/TF (D/TF)		4			
8 [2]	Discharges/TF (D/TF)		2			
9 [523]	Normal	481	14		28	

Table 7.9: Comparison of hypothesis with Dörnenburg Ratios (correspond to Figure 7.15)

		Interpretation by Dörnenburg Ratios					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I [a]**	**N/A [b]**
1 [52]	Discharges (D)		52				
2 [19]	Discharges (D)		17				2
3 [26]	Discharges (D)		12				14
4 [36]	Discharges (D)		5				31
5 [19]	Discharges (D)		1			5	13
6 [29]	Thermal fault (TF)				29		
7 [22]	Thermal fault (TF)				15	5	2
8 [31]	Thermal fault (TF)	4			1	13	13
9 [3]	Discharges/TF (D/TF)		1				2
10 [2]	Discharges/TF (D/TF)		1			1	
11 [516]	Normal	442					74

a. N/I: No interpretation.
b. N/A: Not applicable.

Table 7.10: Comparison of hypothesis with Rogers Ratios (correspond to Figure 7.15)

		Interpretation by Rogers Ratios				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I ***
1 [52]	Discharges (D)		52			
2 [19]	Discharges (D)		19			
3 [26]	Discharges (D)		23			3
4 [36]	Discharges (D)		33			3
5 [19]	Discharges (D)		12			7
6 [29]	Thermal fault (TF)				3	26
7 [22]	Thermal fault (TF)				14	8
8 [31]	Thermal fault (TF)	4	1		14	12
9 [3]	Discharges/TF (D/TF)		3			
10 [2]	Discharges/TF (D/TF)					2
11 [516]	Normal	480	13		16	7

* N/I: No interpretation.

Table 7.11: Comparison of hypothesis with IEC-1999 Ratios (correspond to Figure 7.15)

		Interpretation by IEC Ratios (1999-edition)				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I ***
1 [52]	Discharges (D)		52			
2 [19]	Discharges (D)		19			
3 [26]	Discharges (D)		26			
4 [36]	Discharges (D)		32			4
5 [19]	Discharges (D)		11			8
6 [29]	Thermal fault (TF)				29	
7 [22]	Thermal fault (TF)				18	4
8 [31]	Thermal fault (TF)	4	2		8	17
9 [3]	Discharges/TF (D/TF)		3			
10 [2]	Discharges/TF (D/TF)		1			1
11 [516]	Normal	480	11		20	5

* N/I: No interpretation.

Table 7.12: Comparison of hypothesis with Duval Triangle (1993-edition) (correspond to Figure 7.15)

		Interpretation by Duval Triangle (1993-edition)				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**D/TF**
1 [52]	Discharges (D)		52			
2 [19]	Discharges (D)		19			
3 [26]	Discharges (D)		26			
4 [36]	Discharges (D)		36			
5 [19]	Discharges (D)		19			
6 [29]	Thermal fault (TF)				29	
7 [22]	Thermal fault (TF)		1		19	2
8 [31]	Thermal fault (TF)	4	9		12	6
9 [3]	Discharges/TF (D/TF)		3			
10 [2]	Discharges/TF (D/TF)		1			1
11 [516]	Normal	480	14		22	

As can be observed from these tables, the hypothesis illustrated in Figures 7.14 and 7.15 compares quite well with the interpretation provided by conventional DGA schemes. In fact, the non-existence of the "partial discharge" (PD) region in the hypothesis has also been confirmed since none of the conventional schemes actually detected it. Nevertheless, several discrepancies of interpretation are identified:

- Most conventional DGA schemes are unable to interpret the outer "fault" regions, even though the occurrence of incipient faults is suspected by these schemes. In contrast, the hypothesis illustrated in Figures 7.14 and 7.15 is able to interpret these outer fault regions, which are based on the outcome of detailed statistical analyses on these regions and due to the fact that these regions form a part of the co-centre clusters that are similar in characteristics.
- Even though the Duval Triangle (1993-edition) is capable of detecting the occurrence of duo-faults (i.e. discharges and thermal fault), only one such case is detected, as shown in Table 7.12. In contrast, the hypothesis illustrated in Figures 7.14 and 7.15 has a region dedicated to the duo-fault situation, which is based on the interception between two major incipient fault regions.
- Most of the conventional DGA interpretation schemes, with the exception of Duval Triangle, are unable to interpret all training vectors, which are in fact actual DGA records from oil samples of power transformers. Consequently, the diagnosis of "no interpretation" (N/I) is provided for some of the training vectors even though the presence of incipient fault is suspected by these schemes. In contrast, the hypothesis illustrated in Figures 7.14 and 7.15 is able to provide interpretation for every possible DGA record, based on the best-matching of the DGA record within the identified incipient fault regions.

In conclusion, the hypothesis illustrated in Figures 7.14 and 7.15 is clearly more beneficial since classification of the health and operating condition of power transformers is always guaranteed and is much more convincing.

7.4.6.2 Validation by using actual fault cases

The hypothetical association of revealed features of optimum SOMs with conditions of power transformers has been validated in the context of conventional DGA interpretation schemes. The main objective of this section is to further investigate the validity of the hypothesis through actual, confirmed circumstances.

Herein, the hypothesis is further validated using two actual fault cases, which are the real DGA histories of two power transformers, as confirmed by transformer experts of the National Grid Company (NGC), UK, to indicate some problems or show faults during their operation. In addition, the DGA history of a power transformer that has been observed to be operating normally was also examined for the validation of the "normal operation" region.

DGA trajectories of the abovementioned power transformers are plotted onto the u-matrix via the best-matching means of each DGA history, which contains a time-sequence of DGA records, with neurons of the optimum SOM; the complete operating history of each power transformer is thereby visible via the trajectory plot. The proposed condition of each power transformer, as discerned from its DGA trajectory plot, is then compared with the actual observation of NGC experts. Herein, the optimum SOM for Configuration 10 of the training data was used for the illustration purposes; similar results are also obtained for other chosen optimum SOMs.

Firstly, the DGA trajectory of Transformer A (with voltage level of 275/132 kV and power rating of 240 MVA) is illustrated in Figure 7.16. As seen from Figure 7.16, the trajectory is observed to be gradually moving deeper into the inner clusters of the "discharges" region. This has been confirmed by experts at the NGC, who attributed this phenomenon to the observed arcing/sparking at the clamping plates of this transformer.

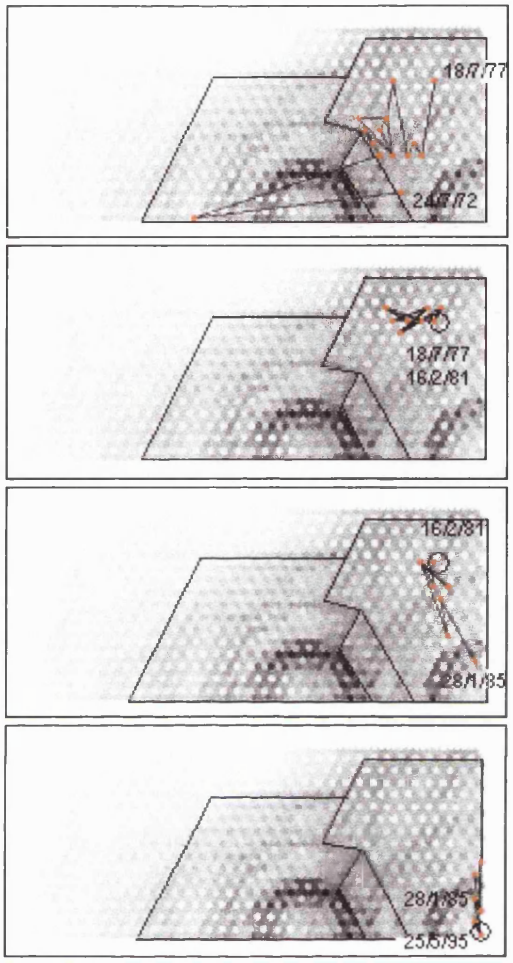


Figure 7.16: The DGA trajectory of Transformer A

Secondly, the DGA trajectory of Transformer B (with voltage level of 275/132 kV and power rating of 240 MVA) is illustrated in Figure 7.17. As seen from Figure 7.17, the trajectory is observed to be gradually moving into inner clusters of the "thermal fault" (TF) region and stays in the inner most cluster from 23/7/89 to 19/5/1993. This has again been confirmed by experts at the NGC, who have attributed the observed phenomenon to a known thermal problem that was peaking in 1991 to 1992.

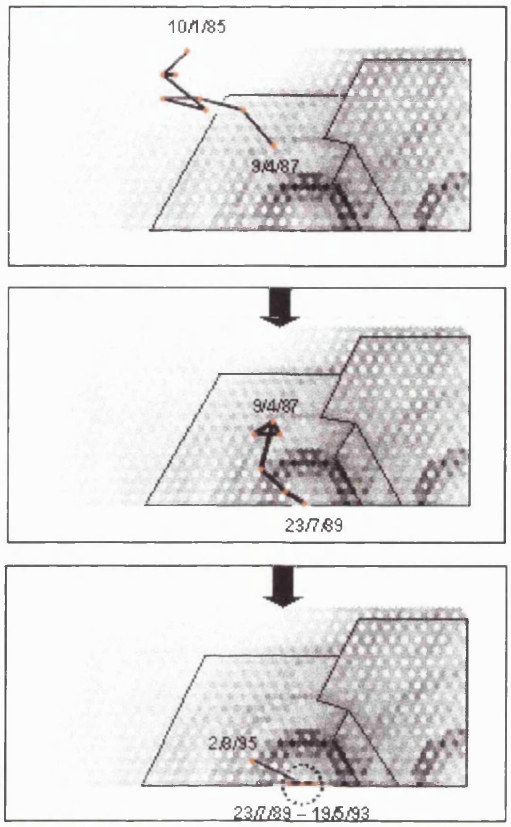


Figure 7.17: The DGA trajectory of Transformer B

Finally, the DGA trajectory of Transformer C (with voltage level of 275/132 kV and power rating of 240 MVA), is illustrated in Figure 7.18. As can be observed from Figure 7.18, the DGA trajectory of this transformer do not venture into any of the identified fault regions, thereby matching the observation of transformer experts at the NGC, who have attributed the variation in the concentrations of key dissolved gases to the normal degradation of insulation oil during normal operating condition.

The foregoing interpretations of the health and operating condition of power transformers, which are based on the hypothesis illustrated in Figure 7.15, correspond very well to the actual observations of transformer experts at NGC. The hypothesis has thus been effectively validated by using actual fault cases. In addition,

the proposed approach has an added advantage of allowing the visualisation of the evolution of operating condition of a power transformer, based on the plotting of its DGA history onto the u-matrix illustration in which various fault regions have been identified.

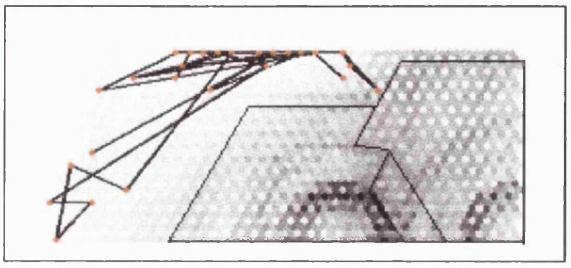


Figure 7.18: The DGA trajectory of Transformer C

7.5 Analysis on the Entire DGA Database of Power Transformers

As reported in the feasibility study, the proposed approach based on the application of SOM has successfully "mined" or unearthed the inherent characteristics of the "Sixsets" DGA data and has presented them in a discernible format. Moreover, the suggested hypothesis based on these revealed features, has been demonstrated to be capable of providing an improved interpretation of the DGA data and an unambiguous visualisation of the DGA history with regard to the health and condition of power transformers.

Herein, the proposed approach is further tested for the analysis of the entire DGA database of 14943 DGA data of power transformers, which is referred to as the "Fullsets" DGA data. The aim of this exercise is to further investigate the capability of the proposed approach for "mining" or unearthing the inherent characteristics of the DGA data that involves a significantly larger amount and variety of samples.

7.5.1 Pre-Processing of the "Fullsets" DGA Data

Similar to the foregoing analysis on the "Sixsets" DGA data, the "Fullsets" DGA data has to be pre-processed before submitting for analysis using the SOM. Similar procedures of pre-processing were applied, as reported in Section 7.3.1.

7.5.2 Data Mining on the "Fullsets" DGA Data

Four sets of DGA data were constructed from the "Fullsets" DGA data. In addition, the "range" scaling method was then applied to these data sets due to its proven usefulness for the effective learning of SOM. Four distinct configurations of training data thus resulted, as shown in Table 7.13.

Table 7.13: The "Fullsets" training data of SOM

Set	Input Components	Configuration of Training Data
A [a]	H_2O, N_2, O_2, CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	1: Range scaling
B [b]	CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	2: Range scaling
C [c]	CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	3: Range scaling
D [d]	H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	4: Range scaling

a. All dissolved gases including moisture.
b. Key dissolved gases from the degradation of cellulose and insulation oil.
c. Dissolved gases that are combustible by nature.
d. Key dissolved gases from the degradation of insulation oil.

Finally, similar configuration and training parameters of SOM were adopted, as previously summarised in Table 7.3. Consequently, four SOMs were trained based on the "Fullsets" DGA data that was configured according to Table 7.13. The

optimum SOM for each configuration of training data was selected according to the procedures outlined in Section 5.5.2.4. However, longer period of training has to be taken due to the size of the "Fullsets" DGA data, which is approximately twenty times that of "Sixsets" DGA data. Besides, significantly less iterations are needed for the training due to the significantly larger size of the "Fullsets" DGA data.

7.5.3 Chosen Optimum SOMs of the "Fullsets" DGA Data

The optimum SOMs of the "Fullsets" DGA data for all configurations of training data can be visualised using the u-matrix and component-plane illustrations. Through observation of the features or patterns as displayed by these optimum SOMs, it is discovered that excellent visualisation of inherent characteristics can be obtained from the optimum SOMs of Configurations 2 and 3 of the training data, and as a consequence, the foregoing optimum SOMs were chosen for further investigation.

The map sizes, number of training iterations and the SOM quality measures for the chosen optimum SOMs are presented in Table 7.14. In addition, Figures 7.19 to 7.22 illustrates the component planes and u-matrices of these optimum SOMs.

Table 7.14: The map sizes, number of training iterations and SOM quality measures for optimum SOMs of the "Fullsets" DGA data

No.	Training Data	Map Size [a] [*x-dim*, *y-dim*]	Number of Training Iterations	SOM Quality Measures [d]	
				AQE [b]	TE [c]
1	Configuration 2	[58 42]	52301 (3½ passes of all training vectors)	0.0289	0.0827
2	Configuration 3	[58 42]	52301 (3½ passes of all training vectors)	0.0237	0.1866

a. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
b. AQE: Average quantisation error
c. TE: Topographic error
d. AQE and TE at the specified number of training iterations during the fine-tuning phase.

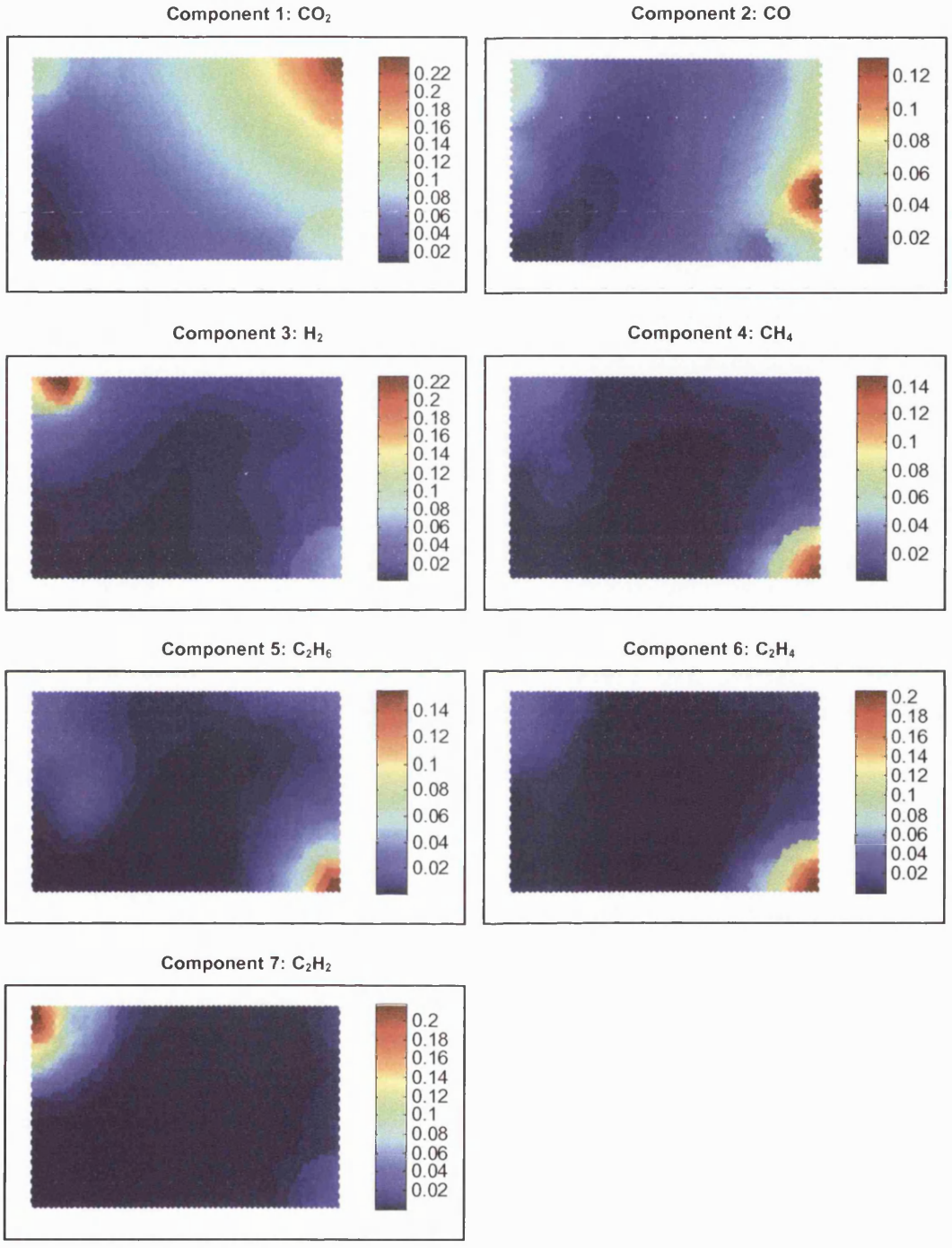


Figure 7.19: Component planes of the optimum SOM for Configuration 2 of the training data (i.e. Set B with "range" scaling)

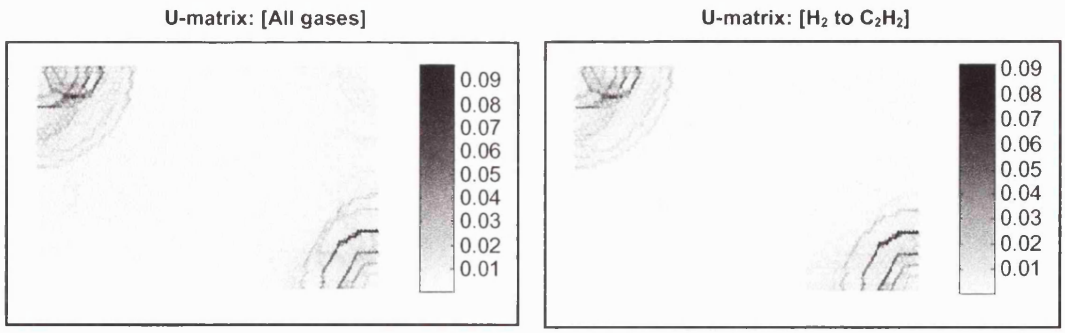


Figure 7.20: U-matrices of the optimum SOM for Configuration 2 of the training data (i.e. Set B with "range" scaling)

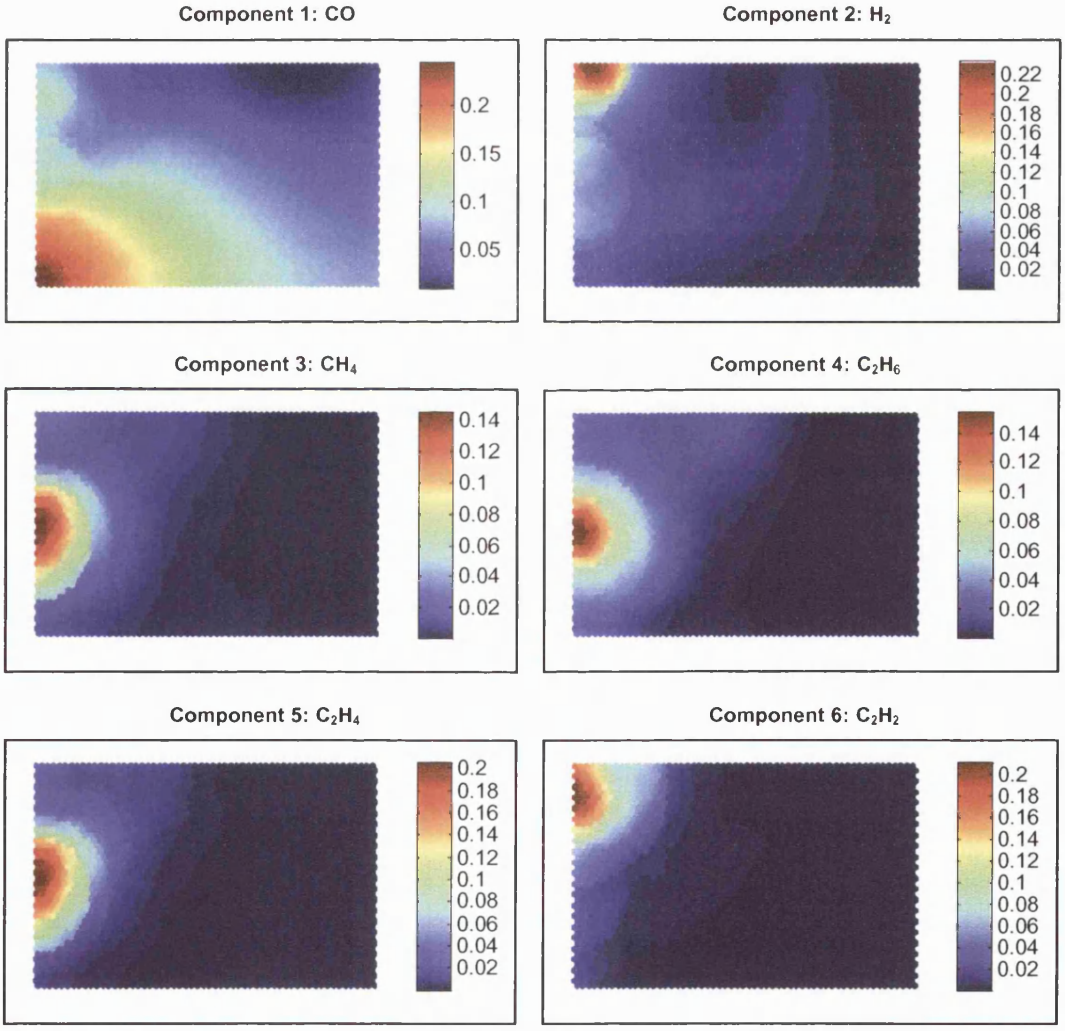


Figure 7.21: Component planes of the optimum SOM for Configuration 3 of the training data (i.e. Set C with "range" scaling")

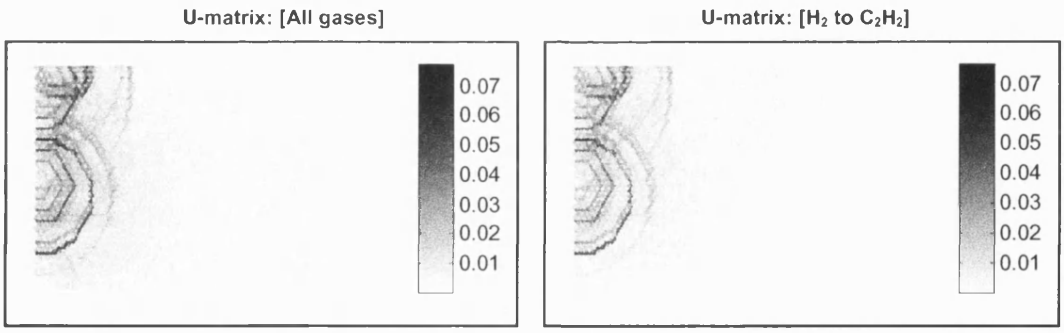


Figure 7.22: U-matrices of the optimum SOM for Configuration 3 of the training data (i.e. Set C with "range" scaling)

Several conclusions can be drawn regarding the foregoing figures of component planes. Firstly, a good correlation can be observed among CH_4, C_2H_6 and C_2H_4, similar to the revealed features of the "Sixsets" DGA data, although, the observed patterns may not look as organised as the "Sixsets" DGA data owing to the larger amount and variety of DGA samples. Secondly, H_2 is only partially correlated with C_2H_2 due to the dissimilarity in the location of revealed patterns. Nevertheless, since the patterns of H_2 and C_2H_2 are very close to each other and the shape of the patterns are similar, it can be concluded that both H_2 and C_2H_2 are related but the relationship may not be as apparent as that of "Sixsets" DGA data.

The observed correlation among CH_4, C_2H_6 and C_2H_4 thus forms a group of apparent co-centre clusters, as observed in u-matrix illustrations of Figure 7.20 and 7.22. Besides, the "partial" correlation of H_2 and C_2H_2 has also led to the formation of a group of clusters (though not co-centre) in the upper-left corner of u-matrix illustrations; interception between H_2 and C_2H_2 clusters is also noticed.

Generally, these two major groups of clusters, i.e. the [CH_4 C_2H_6 C_2H_4] clusters and the [H_2 C_2H_2] clusters, do not overlap or intercept with one another. Notice that a slight interception between these major groups of clusters might be perceived in Figure 7.22. However, on closer examination, this perceived interception or overlapping is actually due to the short proximity between these two groups of

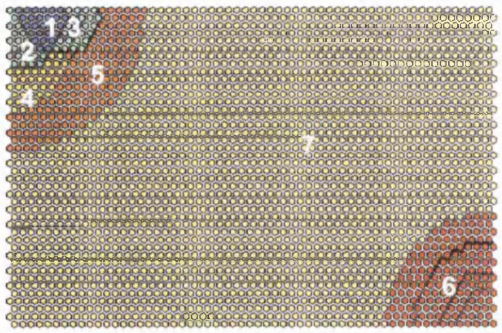
clusters. Furthermore, the fact that no interception is revealed on u-matrices of optimum SOMs for other configurations of training data also confirms this finding. However, in the case of the "Sixsets" DGA data, actual interception between two major groups of co-centre clusters is confirmed due to the fact that identical observation was found for all configurations of the training data.

7.5.4 Analysis on Revealed Features of the "Fullsets" DGA Data

Further statistical analysis has been conducted on the optimum SOMs of Configurations 2 and 3 of the training data. The objective of the foregoing is to examine the distribution and composition of key dissolved gases, i.e. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, across various regions. Essentially, similar steps of analysis were adopted, as previously explained in Section 7.4.4. Results of the analyses are illustrated in Figures 7.23 and 7.24.

Notice that a different approach of presentation was adopted, in which the average concentration of key dissolved gases is shown according to various colour-coded regions, instead of according to each dissolved gas as in the case of "Sixsets" DGA data. The rationale behind this approach is to illustrate the composition of key dissolved gases and to identify dominant dissolved gases for each region. In addition, detail differentiation of clusters for the co-centre clusters of [CH_4 C_2H_6 C_2H_4] was not carried out since it has been proven, as presented in Section 7.4.4, that such clusters in a group of co-centre clusters are indeed similar in characteristics and the average concentration of dominant dissolved gases in these clusters actually increases from the outer clusters to the inner clusters.

Several important conclusions can be drawn from Figures 7.23 and 7.24. Firstly, a similar distribution pattern of average concentration of key dissolved gases is observed for Regions 1 and 3, in which H_2 is the dominant dissolved gas. However, on more detail inspection on those training vectors which correspond to Regions 1 and 3, it is found that there are approximately 75% of training vectors of which the concentration of H_2 is extremely high while other key dissolved gases are low in



Colour coding:

1. 1, 2, 3, 4, 5: [H_2 C_2H_2] clusters
2. 6: [CH_4 C_2H_6 C_2H_4] clusters
3. 7: Apparent cluster not identified

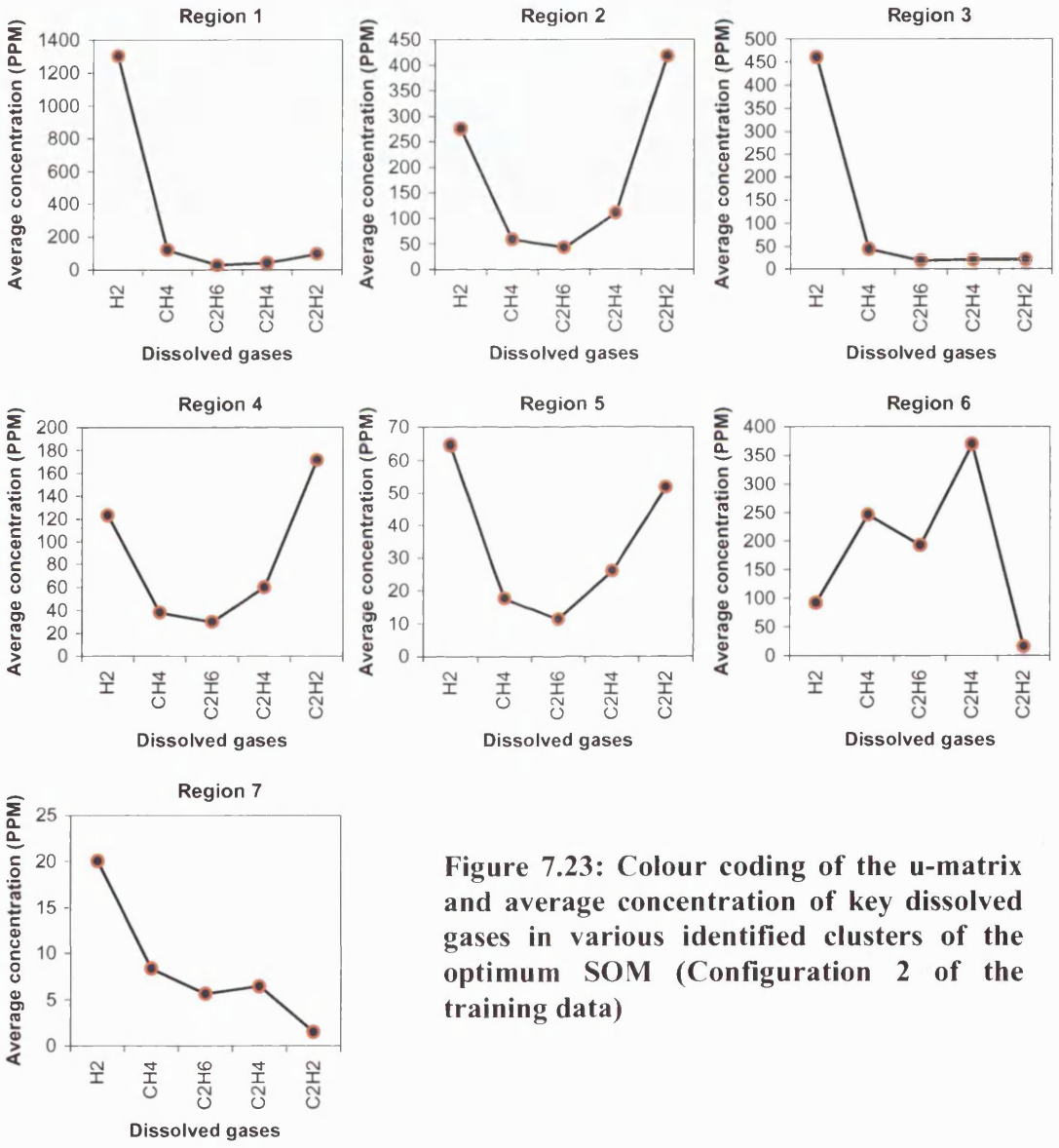
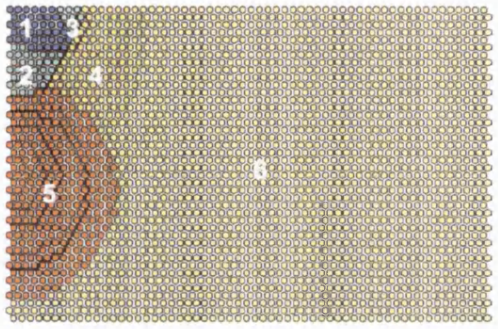


Figure 7.23: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 2 of the training data)



Colour coding:

1. 1, 2, 3, 4: [H_2 C_2H_2] clusters
2. 5: [CH_4 C_2H_6 C_2H_4] clusters
3. 6: Apparent cluster not identified

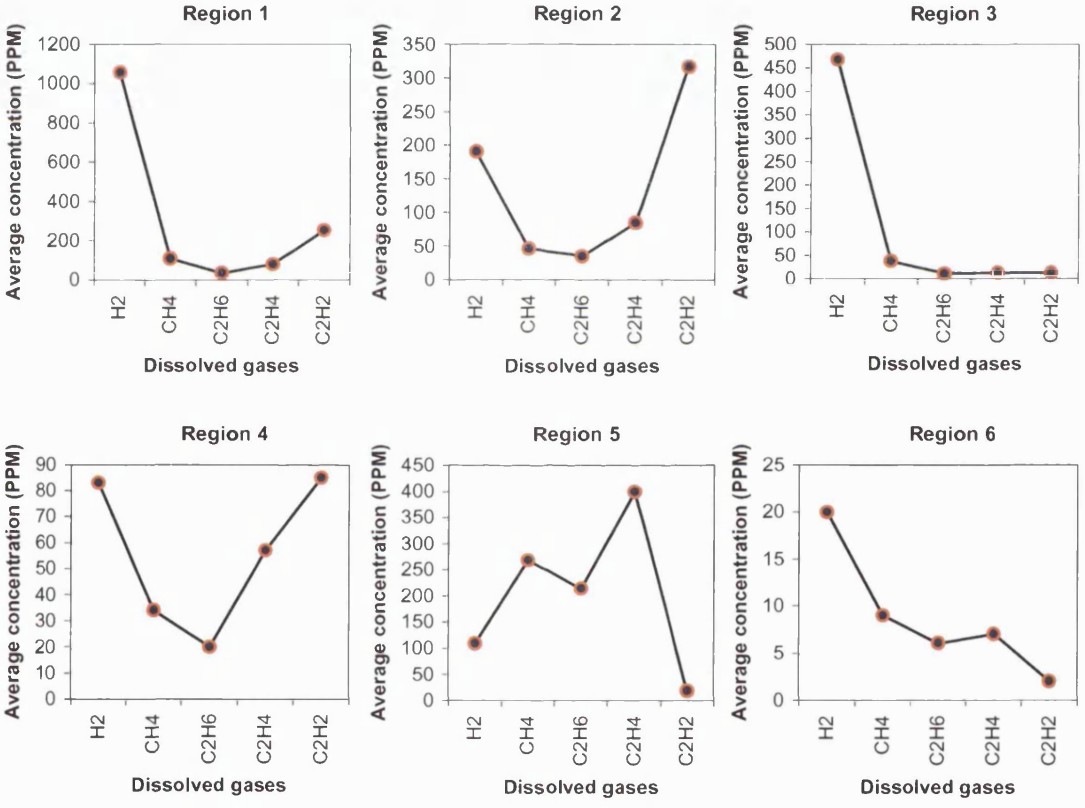


Figure 7.24: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 3 of the training data)

concentration (especially C_2H_2 of which has a concentration close to 0 PPM). However, there are approximately 25% of training vectors of which both H_2 and C_2H_2 is found to be the dominant dissolved gases. Therefore, the two subsets described above actually lead to the dominance of H_2 as seen in Regions 1 and 3.

Secondly, both H_2 and C_2H_2 are found to be the dominant dissolved gases in Regions 2, 4 and 5 of Figure 7.23 and Regions 2 and 4 of Figure 7.24. Thirdly, CH_4, C_2H_6 and C_2H_4 are found to be the dominant dissolved gases in Region 6 of Figure 7.23 and Region 5 of Figure 7.24. Finally, Region 7 of Figure 7.23 and Region 6 of Figure 7.24 are found to contain very low average concentrations of key dissolved gases, since no apparent clusters were formed by the five key dissolved gases in these regions.

In conclusion, one extra feature has been unearthed when the entire DGA database of power transformers was analysed using the SOM, i.e. the discovery of "partial discharge" (PD) pattern due to the extremely high concentration of H_2 in Regions 1 and 3.

7.5.5 Hypothetical Association of Revealed Features

Similar to the "Sixsets" DGA data, an association can be actually established between revealed features of the "Fullsets" DGA data and several conditions of power transformers based on the a priori knowledge on the formation of dissolved gases during the onset and evolution of incipient faults and the previous statistical analyses on chosen optimum SOMs of the "Fullsets" DGA data. An example of the hypothesis is illustrated in Figure 7.25, which is based on the optimum SOM of Configuration 2 of the training data.

As can be observed from Figure 7.25, the majority of the area is dominated by the "normal operating condition" region due to the fact the majority of the "Fullsets" DGA data actually resembles those circumstances of which power transformers are operating faultlessly. In addition, the "partial discharges" (PD) region is dominated

with high concentration of H_2 while the "discharges" region is dominated by high concentrations of H_2 and C_2H_2; both discharges and PD regions can be categorised as being the electrical fault (EF) region. In contrast, the "thermal fault" (TF) region is dominated by high concentrations of CH_4, C_2H_6 and C_2H_4.

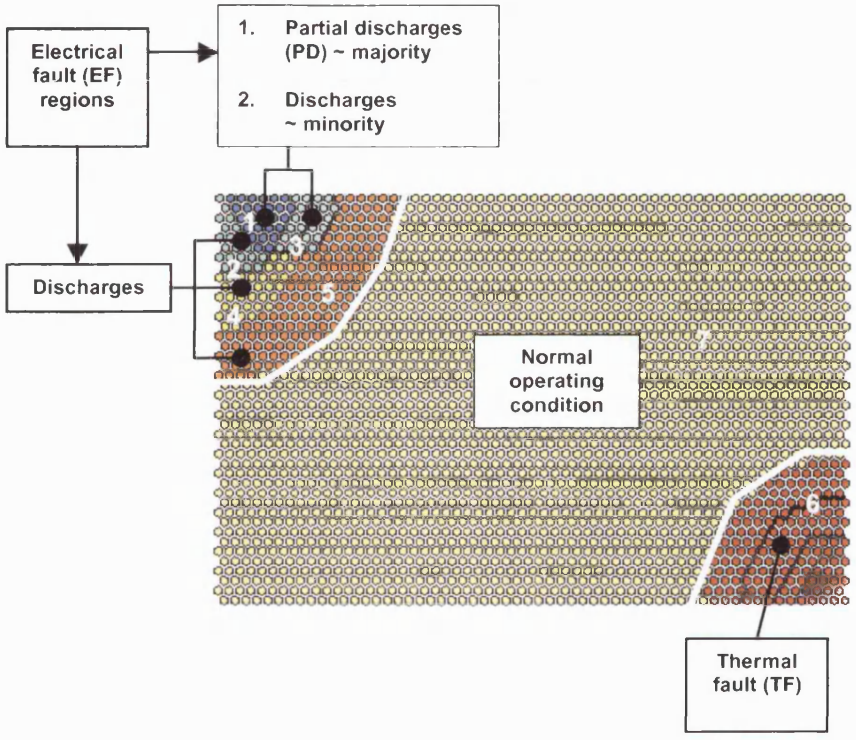


Figure 7.25: Association of revealed features with conditions of power transformers (based on Configuration 2 of the training data)

7.5.6 Validation of the Hypothesis

Again, the hypothetical association of revealed features illustrated in Figure 7.25 has been validated by comparing with several established conventional DGA interpretation schemes and observations by transformer experts at NGC on several actual fault cases, as explained in the following sections:

7.5.6.1 Comparison with conventional DGA interpretation schemes

Those training vectors that correspond to each colour-coded region illustrated in Figure 7.23 were gathered via the best-matching means and subsequently submitted for interpretation by several conventional DGA interpretation schemes after being scaled back to original values. Results of these interpretations are then compared with interpretations as provided by the hypothesis illustrated in Figure 7.25.

Similar to the validation of the "Sixsets" DGA data, four conventional DGA interpretation schemes were utilised, i.e. Dörnenburg Ratios, Rogers Ratios, IEC-1999 Ratios and Duval Triangle (1993-edition). In addition, four main operating conditions of power transformers are to be identified, i.e. normal operation, discharges, partial discharges (PD) and thermal fault (TF). Results of the comparison are shown in Tables 7.15 to 7.18.

Table 7.15: Comparison of hypothesis with Dörnenburg Ratios (correspond to Figure 7.25)

		Interpretation by Dörnenburg Ratios					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I [a]**	**N/A [b]**
1 [193]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		31			22	140
2 [372]	Discharges (D)		262			19	91
3 [78]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		10			6	62
4 [250]	Discharges (D)		22			33	195
5 [375]	Discharges (D)		4			5	366
6 [821]	Thermal fault (TF)				483	83	255
7 [12751]	Normal	10703			16	3	2029

a. N/I: No interpretation.
b. N/A: Not applicable.

Table 7.16: Comparison of hypothesis with Rogers Ratios (correspond to Figure 7.25)

		Interpretation by Rogers Ratios					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I [a]**	**UDF [b]**
1 [193]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		29	102		62	
2 [372]	Discharges (D)		281			91	
3 [78]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		11	33	5	29	
4 [250]	Discharges (D)		163			87	
5 [375]	Discharges (D)		275		2	98	
6 [821]	Thermal fault (TF)				366	453	2
7 [12751]	Normal	11619	203	53	470	402	4

a. N/I: No interpretation.
b. UDF: Undefined, due to the presence of "NaN(s)" in the code(s) (see Section 4.4).

Table 7.17: Comparison of hypothesis with IEC-1999 Ratios (correspond to Figure 7.25)

		Interpretation by IEC-1999 Ratios					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**N/I [a]**	**UDF [b]**
1 [193]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		32	68	73	19	1
2 [372]	Discharges (D)		325		28	19	
3 [78]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		11	5	41	21	
4 [250]	Discharges (D)		202		19	29	
5 [375]	Discharges (D)		291		29	55	
6 [821]	Thermal fault (TF)			1	759	61	
7 [12751]	Normal	11619	189	6	701	234	2

a. N/I: No interpretation.
b. UDF: Undefined, due to the presence of "NaN(s)" in the ratio(s) (see Section 4.6).

Table 7.18: Comparison of hypothesis with Duval Triangle (correspond to Figure 7.25)

		Interpretation by Duval Triangle (1993-edition)					
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**D**	**PD**	**TF**	**D/TF**	**UDF ***
1 [193]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		52	45	94	1	1
2 [372]	Discharges (D)		372				
3 [78]	1. Partial discharges (PD) ~ majority 2. Discharges ~minority		16	1	51	10	
4 [250]	Discharges (D)		250				
5 [375]	Discharges (D)		368		2	5	
6 [821]	Thermal fault (TF)		18		747	54	2
7 [12751]	Normal	11619	308	6	753	65	

* UDF: Undefined, due to the presence of "NaN(s)" in the %gas(es) (see Section 4.7).

As can be observed from Tables 7.15 to 7.18, various discrepancies in the interpretation are identified, as described below:

- Dörnenburg Ratios is not capable of identifying those training vectors that contain the PD pattern discovered in the foregoing statistical analysis. Nevertheless, it did correctly interpret some of the "discharges" and TF patterns. In addition, constrains imposed by Dörnenburg for the applicability of his method also lead to a large amount of DGA data which is not possible to be interpreted (i.e. "not applicable"). Finally, Dörnenburg Ratios is also not able to provide interpretation for some of the training vectors that satisfied his constrains but not covered by the combinations of ratios.
- Interpretations provided by Rogers Ratios correspond very closely to the hypothesis, in which it has identified PD patterns in Regions 1 and 3. Moreover, it has correctly identified the "discharges" and TF patterns.

However, it is unable to provide interpretation for some of the training vectors due to its inability to cover every possible combination of ratios.

- Although the IEC-1999 Ratios technique has identified some PD patterns in Regions 1 and 3, it also incorrectly classified some of the supposedly PD patterns as TF. Owing to the extremely high concentration of H_2 (over 1000 PPM) in Regions 1 and 3 and relatively low concentrations of other key dissolved gases, the majority of training vectors should therefore indicate the PD phenomenon rather than TF. Moreover, the identification of TF patterns in Regions 1 and 3 is also not convincing since CH_4, C_2H_6 and C_2H_4 are not the dominant dissolved gases in these regions. Nevertheless, the IEC-1999 Ratios technique has correctly identified most of the "discharges" and TF patterns in Regions 2, 4, 5 and 6. Finally, the IEC-1999 is unable to provide interpretation for some of the training vectors owing to its inability to cover every possible combination of ratios.
- Duval Triangle has correctly identified the "discharges" and TF patterns in Regions 2, 4, 5 and 6. However, it did suffer from the same weakness as the IEC-1999 Ratios technique since it has incorrectly interpreted some of the PD patterns as TF.

When compared with conventional DGA interpretation schemes, interpretations provided by the hypothesis illustrated in Figure 7.25 are more convincing, due to the fact that the association of regions with health and operating conditions of power transformers is established based on the a priori knowledge on the formation of dissolved gases during the onset and evolution of incipient faults and detailed statistical analysis on identified clusters of key dissolved gases, as presented in Section 7.5.4.

In addition, the hypothesis illustrated in Figure 7.25 is able to provide plausible interpretation for every DGA data based on the best-matching of each DGA data with reference vectors of SOM neurons within each region.

7.5.6.2 Validation by using actual fault cases

The performance of the hypothesis illustrated in Figure 7.25 compares quite favourably with several conventional DGA interpretation schemes, as presented in Section 7.5.6.1. Herein, the hypothesis is further validated using several actual, observed fault cases that have been confirmed by transformer experts at NGC. Similar procedures of validation to that of "Sixsets" DGA data were utilised. Results of the validation are described sequentially in this section.

Firstly, the DGA trajectory of Transformer A (with voltage level of 275/132 kV and power rating of 240 MVA) is illustrated in Figure 7.26. It appears as if can be observed, the DGA trajectory gradually moving deeper into the "PD/discharges" region, arriving at the centre of this region on date 6/4/94. This observation has been confirmed by experts at NGC, who indicated that this particular transformer actually failed in 1994.

Secondly, the DGA trajectory of Transformer B (with voltage level of 275/132 kV and power rating of 240 MVA) is illustrated in Figure 7.27. As can be observed from this figure, the DGA trajectory always fluctuates within the inner "discharges" region from 24/7/72 to 5/6/96. This has been confirmed by experts at the NGC, in which this phenomenon was attributed to the observed arcing/sparking at the clamping plates of this transformer.

Thirdly, the DGA trajectory of Transformer C (with voltage level of 275/132 kV and power rating of 240 MVA) is illustrated in Figure 7.28. As can be observed from this figure, the trajectory appears to be gradually moving into the inner clusters of the "thermal fault" (TF) region and stays in inner clusters from 23/7/89 to 19/5/1993. This has been confirmed by the experts at the NGC, who attributed this observed phenomenon to a known thermal problem that was peaking in 1991 to 1992.

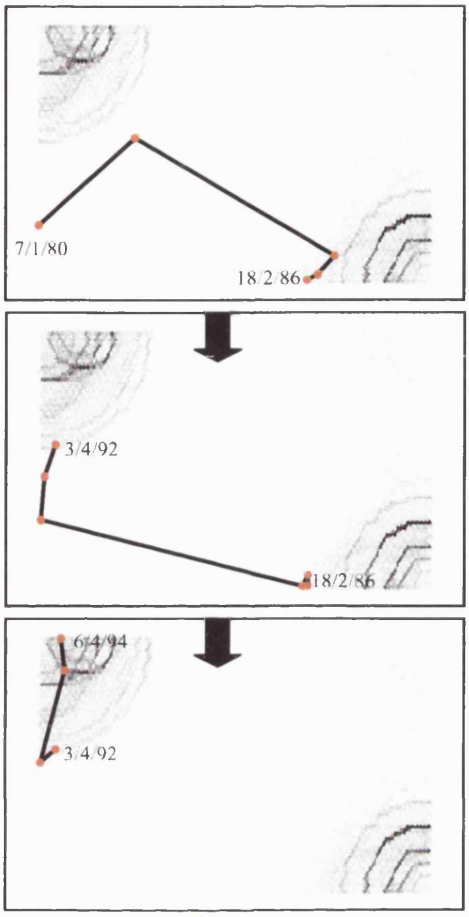


Figure 7.26: The DGA trajectory of Transformer A

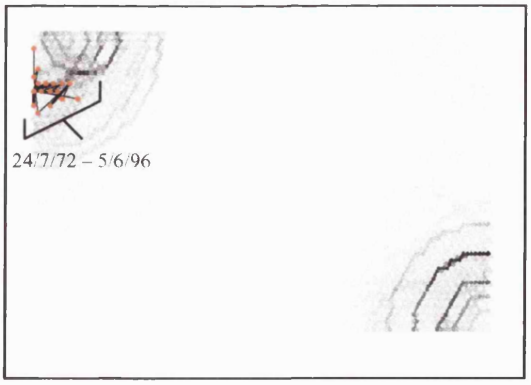


Figure 7.27: The DGA trajectory of Transformer B

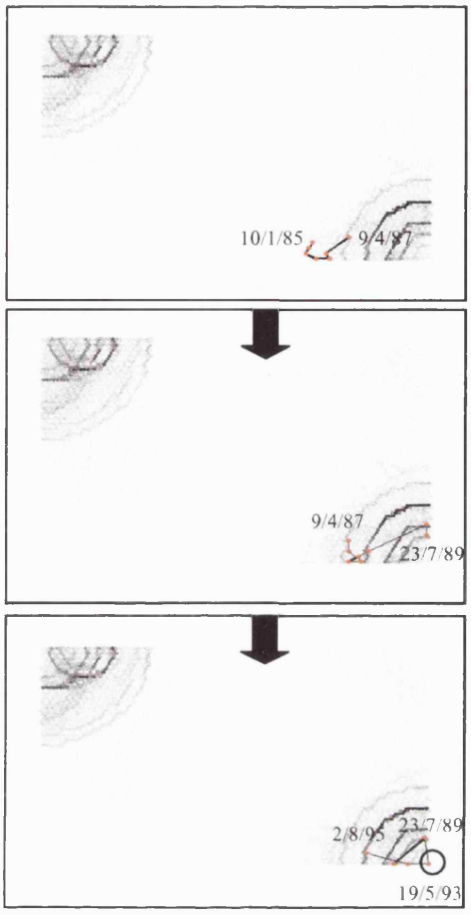


Figure 7.28: The DGA trajectory of Transformer C

Finally, the DGA trajectory of a normal operating transformer, i.e. Transformer D (with voltage level of 275/132 kV and power rating of 240 MVA), is illustrated in Figure 7.29. As can be observed, the DGA trajectory of this transformer does not venture into any of the identified fault regions, thereby corresponding very well with the observation of transformer experts at the NGC.

In conclusion, the foregoing interpretations of the health and operating condition of power transformers, which are based on the hypothesis illustrated in Figure 7.25, correspond very well with actual observations of transformer experts at NGC. Furthermore, the proposed approach has an added advantage of allowing the

visualisation of the evolution of operating condition of a power transformer, based on the plotting of its DGA history onto the u-matrix illustration.

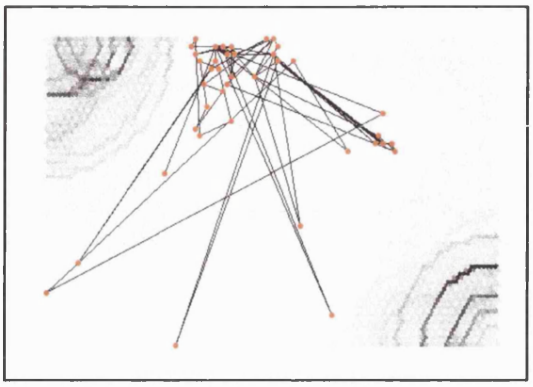


Figure 7.29: The DGA trajectory of Transformer D

7.6 Advantages of the Proposed Approach

The proposed approach based on the SOM has the following advantages if compared to the conventional DGA interpretation schemes and other artificial intelligence (AI) approaches:

- The modelling process of the proposed approach does not depend on conventional DGA interpretation schemes or actual fault cases; the proposed approach is based entirely on routine DGA samples which exist in abundance and are easy to obtain from database. Conventional DGA schemes and actual fault cases are simply needed for the validation of the proposed approach once modelled.
- The interpretation as provided by the proposed approach is more convincing since it is based on common-sense knowledge about the formation of dissolved gases during the onset and evolution of incipient faults and more importantly, based on detailed statistical analyses on the revealed features of the optimum SOMs of the DGA data.

- Interpretation on the health and operating condition of a power transformer is always guaranteed since the proposed approach is able to provide reliable interpretation for every possible combination of dissolved gases.
- The uncertainty and ambiguity in interpretation, as encountered in conventional DGA schemes, have been eliminated since the interpretation as provided by the proposed approach is founded on inherent characteristics of the actual DGA data.
- Visualisation on health and condition of any power transformer is now a reality via the powerful means of the trajectory plotting of its DGA history on the u-matrix illustration where various fault regions have been identified. Hence, the onset and evolution of incipient fault can be visualised effectively. Moreover, early detection of abnormality via this means could prevent catastrophic failure of the power transformer.
- The proposed approach is cost-effective to implement since it only based on routine DGA samples, once a software has been implemented to automate the process of SOM training and selection of optimum SOMs.
- The proposed approach, whilst not aiming to replace entirely the dependence on widely acknowledged conventional DGA interpretation schemes and valuable knowledge of human experts, can be used as a decision support tool for transformer engineers who wish to monitor the health and operating condition of power transformers.

7.7 Recommended Implementation of the Proposed Approach

The proposed approach, which is based on the application of a DM method known as the SOM, has been shown to be capable of providing an improved analysis and interpretation of the DGA data of power transformers. The advantages of the proposed approach have been presented in the previous section.

The proposed approach is generic and can be applied to the DGA data of power transformers from other power utilities, on condition that a set of DGA data, which

consists of DGA samples of power transformers of similar type and functionality, is collected. In essence, the DGA data can be from a family of power transformers of identical manufacturers, voltage levels and power ratings, or a mixture of these. In all cases, the proposed approach has been shown to be capable of learning the inherent characteristics of the DGA data, as previously demonstrated for the "Sixsets" and "Fullsets" DGA data.

Once the association of the revealed features with the health and operating condition of power transformers have been established, an individual power utility could visualise and monitor the evolution of the operating condition of a particular power transformer based on the trajectory plotting of its DGA history. Therefore, the movement of the trajectory into any of the determined fault regions can be reported swiftly and corrective actions could be taken. The approach is flexible and it is up to an individual power utility to decide when an alarm should be generated based on the movement of the DGA trajectory.

Basically, the approach can be utilised in two different manners. Firstly, a DGA sample of a particular power transformer at a particular date of sampling can be mapped onto the interpreted map; the region in which the best matching unit (BMU) of this DGA sample lies defines the condition or the state of health of this particular power transformer, as illustrated in Figure 7.30.

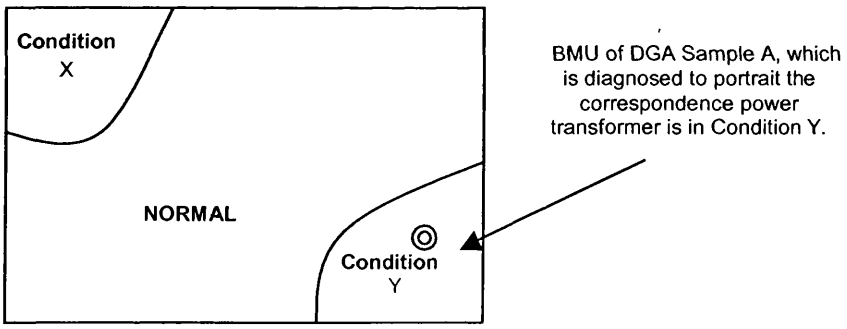


Figure 7.30: Fault diagnosis based on the proposed approach

Secondly, the trajectory of DGA history for a particular transformer can be plotted onto the interpreted map again via the best-matching means; the evolution of the state of health or operating condition of the power transformer can be visualised. Essentially, the gradual movement of the trajectory deeper into any of determined fault region should warrant the attention of transformer engineers since it indicates that the concentrations of key dissolved gases are on the increase, and hence the suspected incipient fault may be persistent. This approach of implementation is illustrated in Figure 7.31.

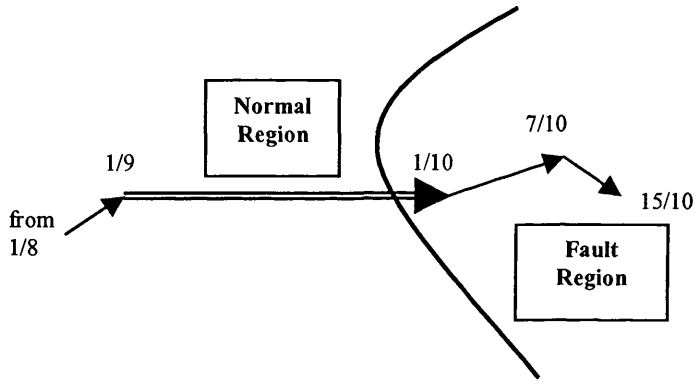


Figure 7.31: Monitoring of the evolution of the operating condition of a power transformer based on the proposed approach

However, it should be noted that the proposed approach is only capable of indicating the health and condition of power transformers, similar to the conventional DGA interpretation schemes. It is able to detect the onset and to show the evolution of any incipient fault but it is not capable of indicating the location of the fault once detected. The detection of the actual cause and location of the incipient fault is only possible through the utilisation of diagnostic tests.

7.8 Summary

The proposed approach, which is based on a DM method known as the SOM, has been applied for the analysis and interpretation of the DGA data of power

transformers. In essence, two sets of DGA data have been examined; the first set corresponds to a total of 755 DGA records extracted from power transformers of various manufacturers, voltage levels and power rating and the second set corresponds to the entire DGA database of 14943 DGA records. In all instances, the proposed approach has demonstrated its ability of providing improved interpretation on the health and condition of power transformers if compared to a number of commonly applied conventional DGA interpretation schemes. Importantly, its performance has been validated with real fault cases, albeit for a limited number of cases. Furthermore, it has been demonstrated that the proposed approach is capable of providing the visualisation of the state of the health and operating condition of power transformers, which is an additional, important advantage and of major benefit to transformer engineers.

CHAPTER 8

DATA MINING ON THE DGA DATA OF TRANSFORMER BUSHINGS USING THE SOM

8.1 Introduction

As reported in Chapter 7, the proposed approach based on the self-organising map (SOM) has been successfully employed to "mine" or unearth the inherent characteristics of the dissolved gas analysis (DGA) data of power transformers; these characteristics have been subsequently associated with several conditions of power transformers. In this chapter, a further study is presented, in which the same developed approach was employed in an attempt for the analysis and interpretation of the DGA data of transformer bushings. Various aspects of the study are presented in the following sections of the chapter.

8.2 General Design of a Transformer Bushing

Essentially, transformer bushings provide a means of electrical connection between overhead line or bus bar to the main unit of power transformer. A power transformer is typically equipped with three low-voltage (LV) and three high-voltage (HV) bushings, which are located on the top of the transformer tank.

The general design of a transformer bushing is illustrated in Figure 8.1, which is accompanied by a brief description of various constituent parts of a transformer bushing.

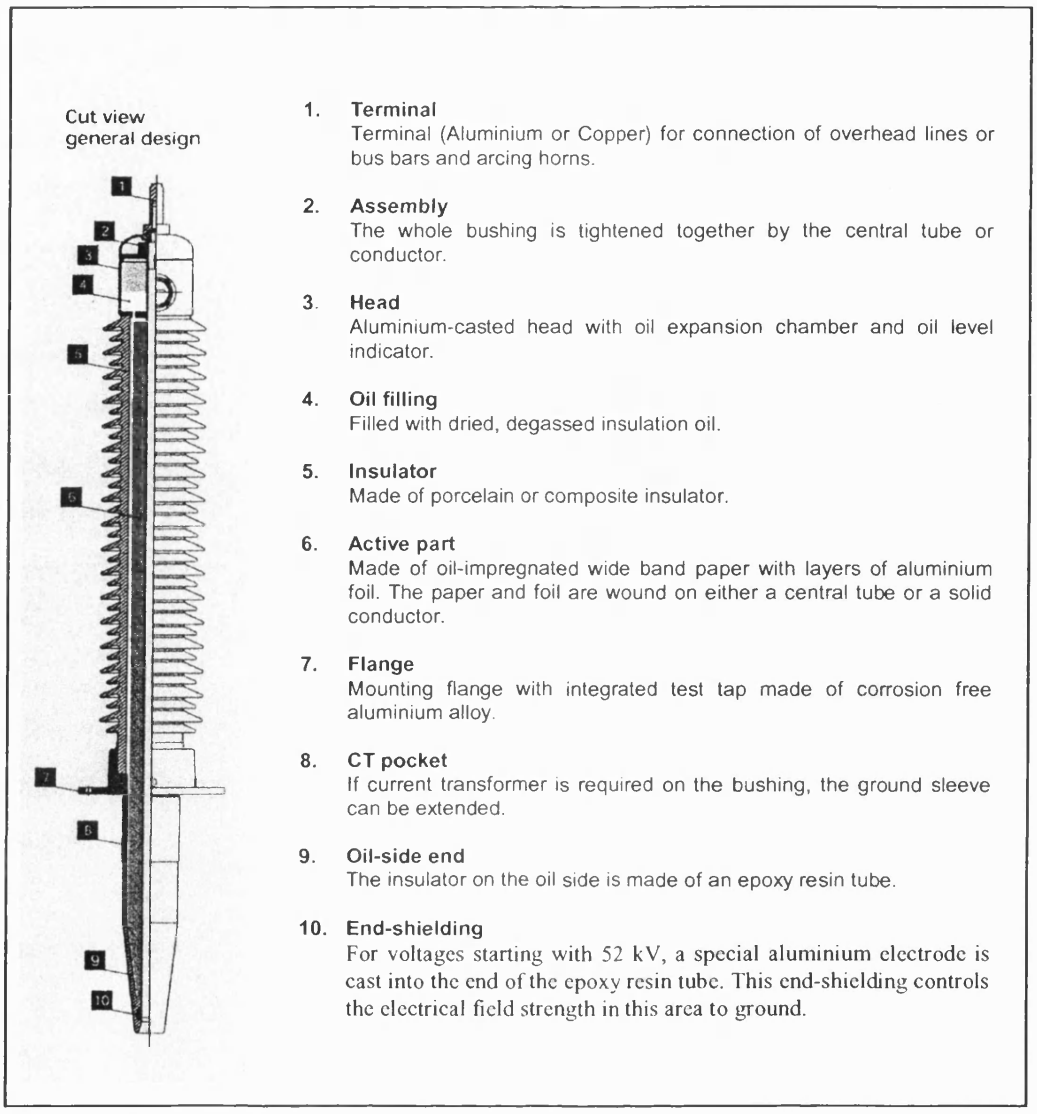


Figure 8.1: General design of a transformer bushing

As seen from Figure 8.1, insulation oil of a transformer bushing is contained within a tightly enclosed environment between the porcelain insulator and the cellulose insulation. In addition, the volume of oil is considerably smaller if compared to that of transformer tanks. Consequently, it can be expected that the inherent characteristics of dissolved gases, which are formed either due to the normal degradation of the oil/paper insulation or due to the onset or evolution of faults, to be quite distinct from that of power transformers.

8.3 Challenges in Interpretation of the DGA Data

Essentially, DGA test is only performed every 12 years on a normally operating transformer bushing owing to the difficulty and high cost of performing oil sampling on bushings. Nevertheless, this sampling interval may be reduced to either three or six years if problems or faults are suspected based on pre-defined levels of certain dissolved gases. Hence, the interval between subsequent DGA tests on transformer bushings is considerably longer when compared with the typically three-monthly interval on transformer tanks. The DGA data of transformer bushings might not thus be as informative and useful as the DGA data of power transformers for condition monitoring and fault diagnosis purposes, owing to the very limited amount of DGA data available.

According to engineers of the NGC, they tend not to be very interested in the identification of the "type" of faults from the DGA data of transformer bushings; the indication of the "presence" of faults is regarded to be more important. In addition, since the cost of replacing a failed transformer bushing is significantly lower compared to the exorbitant cost of repairing or even replacing a failed main unit of power transformer, once a test has been performed two options are generally available: replace the bushing or to leave it as it is. Moreover, engineers are very restricted in re-sampling of oil due to the small oil-volume and further investigation is virtually impossible due to the high cost and inconvenience of repeating the oil sampling. Owing to the foregoing reasons, DGA data of transformer bushings are typically used by engineers for the "indication" or "detection" of abnormalities or faults in transformer bushings rather than for the diagnosis of a suspected fault and monitoring of its evolution.

Despite the limited usefulness and practicality of the DGA data of transformer bushings, IEC 60599 Standard [20] has recommended a means of relating the DGA data of bushings (not specifically transformer bushings) with several common incipient faults of bushings, and this is remarkably similar to the interpretation for

the DGA data of power transformers. However, this approach is not adopted by engineers of the NGC for reasons described previously.

In spite of all the foregoing challenges, an attempt was made to apply the proposed approach for the analysis of the DGA data of transformer bushings. It was hoped that interesting features could be unearthed from the data and consequently association could possibly be established between revealed features and the health and conditions of transformer bushings. Therefore, the investigation as reported in this chapter aims at presenting a systematic analysis of the DGA data of transformer bushings using the SOM so as to ascertain the possibility of proposing a potentially useful and improved approach for the interpretation of the DGA data of transformer bushings.

8.4 DGA Data of Transformer Bushings

The DGA data of LV and HV bushings of power transformers was obtained courtesy of the NGC, UK. This set of DGA data, which comprises of 2008 samples, corresponds to transformer bushings of various phases (i.e. Red (A), Yellow (B) or Blue (C) phase), as illustrated in Table 8.1. The date of sampling of these DGA data ranges from 5/5/1984 to 13/3/2001.

Table 8.1: The DGA data of transformer bushings

Subset	Description	Number of Data
1	A/Red LV Bushings	172
2	B/Yellow LV Bushings	174
3	C/Blue LV Bushings	175
4	A/Red HV Bushings	516
5	B/Yellow HV Bushings	488
6	C/Blue HV Bushings	483

In addition, the format of the DGA data of transformer bushings is illustrated in Table 8.2 below:

Table 8.2: Format of the DGA data of transformer bushings

A	C	D	E	F	G	H	I
Plant_ID	**Point-Code**	**NGC Number**	**Amis**	**Type**	**Sampdate**	**Gas**	**CO**
ABTH2_SGT1	12	T3998	146	Bush_sgt	15/08/96	7.61	129
ABTH2_SGT1	13	T3998	146	Bush_sgt	15/08/96	5.57	132
ABTH2_SGT1	14	T3998	146	Bush_sgt	15/08/96	6.66	114

J	K	L	M	N	O	P	Q	R
CH_4	**CO_2**	**C_2H_4**	**C_2H_6**	**C_2H_2**	**H_2**	**O_2**	**N_2**	**H_2O**
63	776	7	80	< 0.2	20	843	77292	50
38	604	3	53	< 0.2	14	977	57309	12
92	776	7	80	< 0.2	20	843	77292	50

As shown in the table, nine dissolved gases and moisture (i.e. H_2O) are recorded in a single DGA data, accompanied by some background information such as the location (Plant_ID) and identity (AMIS and NGC Number) of the corresponding power transformer of which the bushing is situated, reference of sampling for each phase (Point_Code), the type of bushing (Type), date of sampling (Sampdate) and the percentage composition of dissolved gases in an oil sample (Gas). Note that all dissolved gases and moisture are recorded in terms of part-per-million (PPM), which represents a certain quantity of a dissolved gas, which is measured in micro-litre, in 1 litre of an oil sample. The DGA data of transformer bushings will be referred to as the "bushings" DGA data in subsequent sections of this chapter.

8.5 Pre-Processing of the "Bushings" DGA Data

The "bushings" DGA data has to be pre-processed before submitted for analysis using the SOM. Firstly, each "blank" (no data entry or record) in the "bushings"

DGA data was replaced with the notation "NaN" (i.e. Not-a-Number) so as to minimise the loss of valuable information even if the concentration of a particular dissolved gas is missing. The SOM is capable of handling these "NaNs" by ignoring those components of reference vectors that contain "NaNs" for the calculation of Euclidean distances during the SOM training process.

Secondly, there are doubts about the credibility of "0 PPM" in atmospheric dissolved gases (i.e. N_2, O_2 and CO_2) and products of cellulose degradation (i.e. CO_2 and CO), due to the fact that it is not possible for the transformer bushings to be completely sealed-off from the outside environment and also due to the significant presence of cellulose insulation within the units. Consequently, each "0 PPM" record was replaced by the notation "NaN" so as to avoid confusion in subsequent analyses. These "NaNs" are handled by SOM in a similar manner as described previously.

Finally, those entries of dissolved gases that are in the format of "$< x$", in which "$< x$" is the "analytical detection limit" of each dissolved gas, were converted to reasonably small values in a systematic manner, as shown in Table 8.3.

Table 8.3: Conversion of "$< x$" entries into assigned values

No.	Dissolved Gas	Original Format (PPM)	Assigned (PPM)
1	N_2	< 50	49
2	O_2	< 50	49
3	CO_2	< 50	49
4	CO	<10	9
5	H_2	< 1	0.9
6	CH_4	< 1	0.9
7	C_2H_6	< 1	0.9
8	C_2H_4	< 1	0.9
9	C_2H_2	< 0.2	0.19

8.6 Data Mining on the "Bushings" DGA Data

The pre-processed DGA data was analysed using the SOM. Four sets of DGA data were constructed from nine dissolved gases and moisture; the "range" scaling method was then applied to each set of DGA data due to its proven usefulness for data analysis. The four configurations of training data are shown in Figure 8.4.

Table 8.4: The "Bushings" training data of SOM

Set	Input Components	Configuration of Training Data
A [a]	H_2O, N_2, O_2, CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**1**: Range scaling
B [b]	CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**2**: Range scaling
C [c]	CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**3**: Range scaling
D [d]	H_2, CH_4, C_2H_6, C_2H_4, C_2H_2	**4**: Range scaling

Note:
a. All dissolved gases including moisture.
b. Key dissolved gases from the degradation of cellulose and insulation oil.
c. Dissolved gases that are combustible by nature.
d. Key dissolved gases from the degradation of insulation oil.

The basis for adopting various combinations of input components is to search for interesting features in the optimum SOM for each configuration of the training data. Besides, it is also the aim to identify those key or important dissolved gases that actually contribute to the observed features or patterns in the optimum SOMs.

Besides, configuration and training parameters as presented in Table 7.3 were adopted for the training of SOM since the technique is generic and identical configuration and training parameters can be utilised. Consequently, four SOMs were trained based on the "bushings" DGA data that was configured according to Table 8.4. Finally, the optimum SOM for each configuration of the training data was selected according to the procedures outlined in Section 5.5.2.4.

8.7 Optimum SOMs of the "Bushings" DGA Data

The optimum SOM for each configuration of the training data is first visualised using the component-plane illustration, as shown in Figures 8.2 to 8.5. In addition, the map sizes, number of training iterations and SOM quality measures for each optimum SOMs are illustrated in Table 8.5.

As can be observed from Figures 8.2 to 8.5, revealed features of the "bushings" DGA data are different from that of the DGA data of power transformers. In fact, the previously observed correlation in [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4] for the DGA data of power transformers is not observed in revealed features of the "bushings" DGA data. On the contrary, the locations and patterns of these gases are similar to one another.

Moreover, with the presence of the following dissolved gases: CO, H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, two apparent patterns are observed in the CH_4 component-plane. The first pattern is correlated with H_2, C_2H_6, C_2H_4 and C_2H_2 while the second pattern corresponds to high concentration of CO. These features are not revealed when CO does not form one of the input components, as can be perceived from Figure 8.5, which shows the component planes of the optimum SOM that are based on key dissolved gases from the degradation of insulation oil only. It can thus be deduced that "key" dissolved gases of the "bushings" DGA data which have contributed to the observed features, are CO, H_2, CH4, C_2H_6, C_2H_4 and C_2H_2.

In addition, revealed features of the "bushings" DGA data can also be visualised using the u-matrix illustration. Two examples of the u-matrix are shown in Figures 8.6 and 8.7, which correspond to the optimum SOMs of Configurations 1 and 2 of the training data. Although these u-matrices are formed using H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, the relevance of CO will be considered when it comes to performing the detailed statistical analysis on revealed features of the "bushings" DGA data, as will be presented in the next section of this chapter.

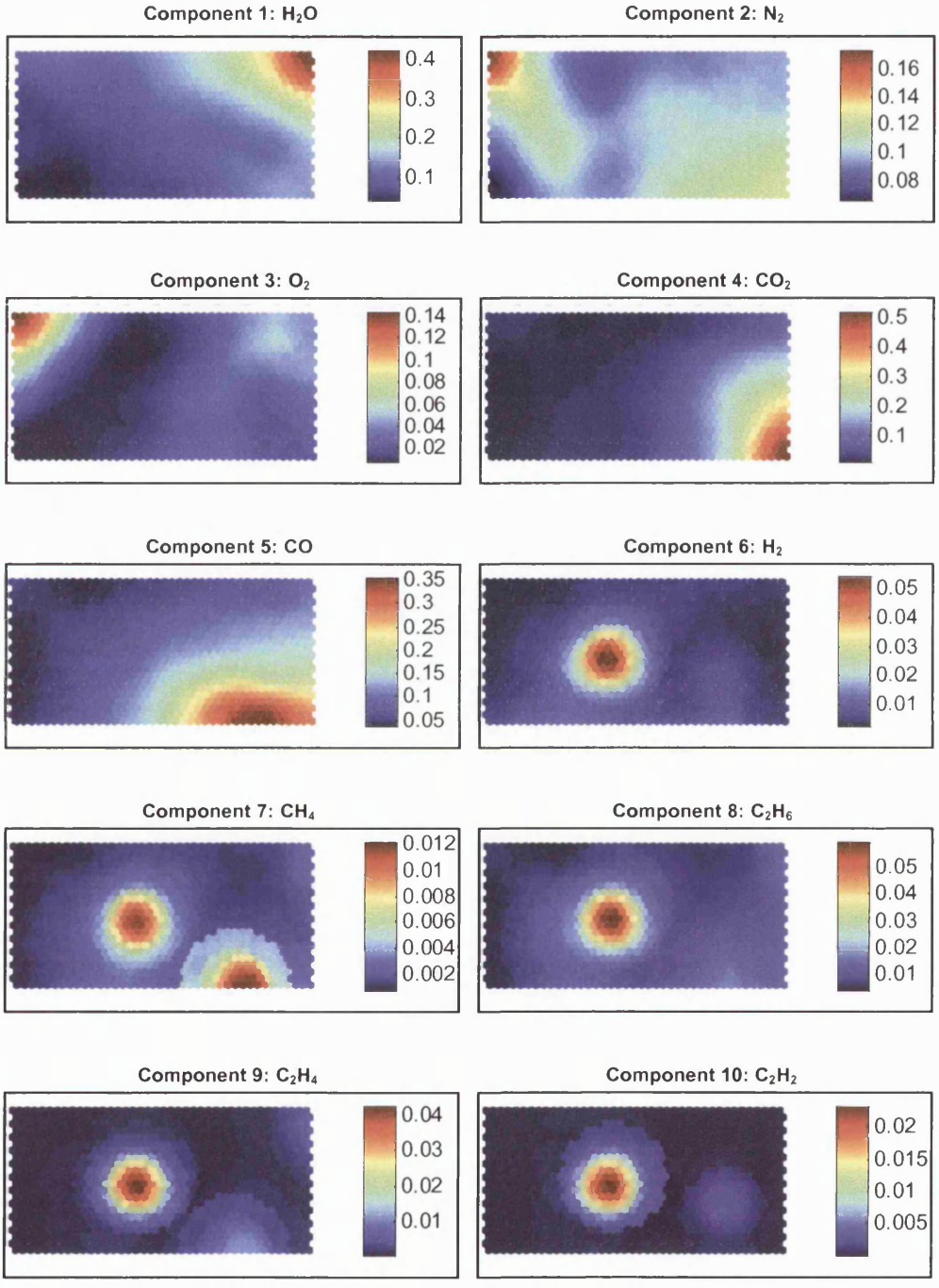


Figure 8.2: Component planes of the optimum SOM for Configuration 1 of the training data (i.e. Set A with "range" scaling)

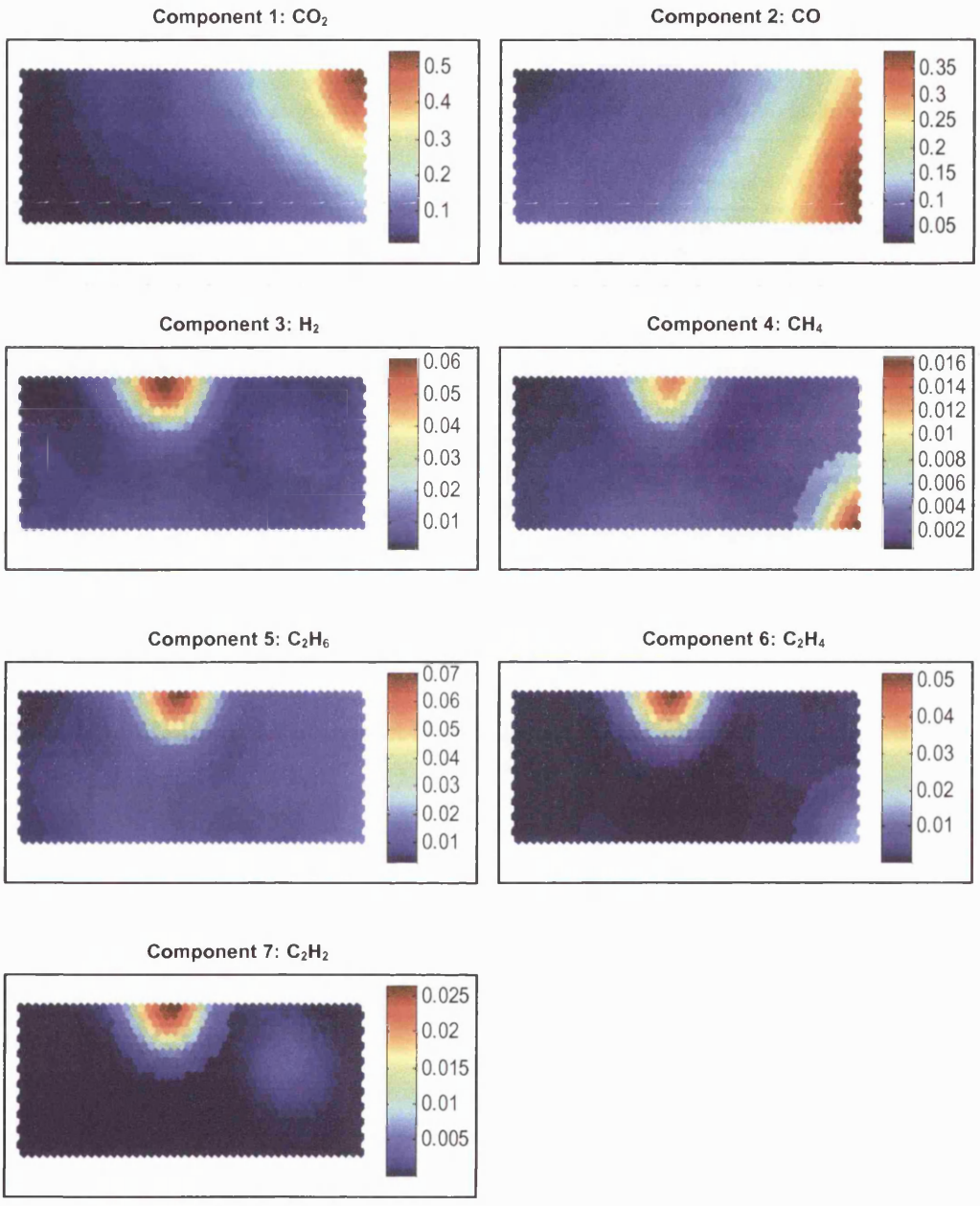


Figure 8.3: Component planes of the optimum SOM for Configuration 2 of the training data (i.e. Set B with "range" scaling)

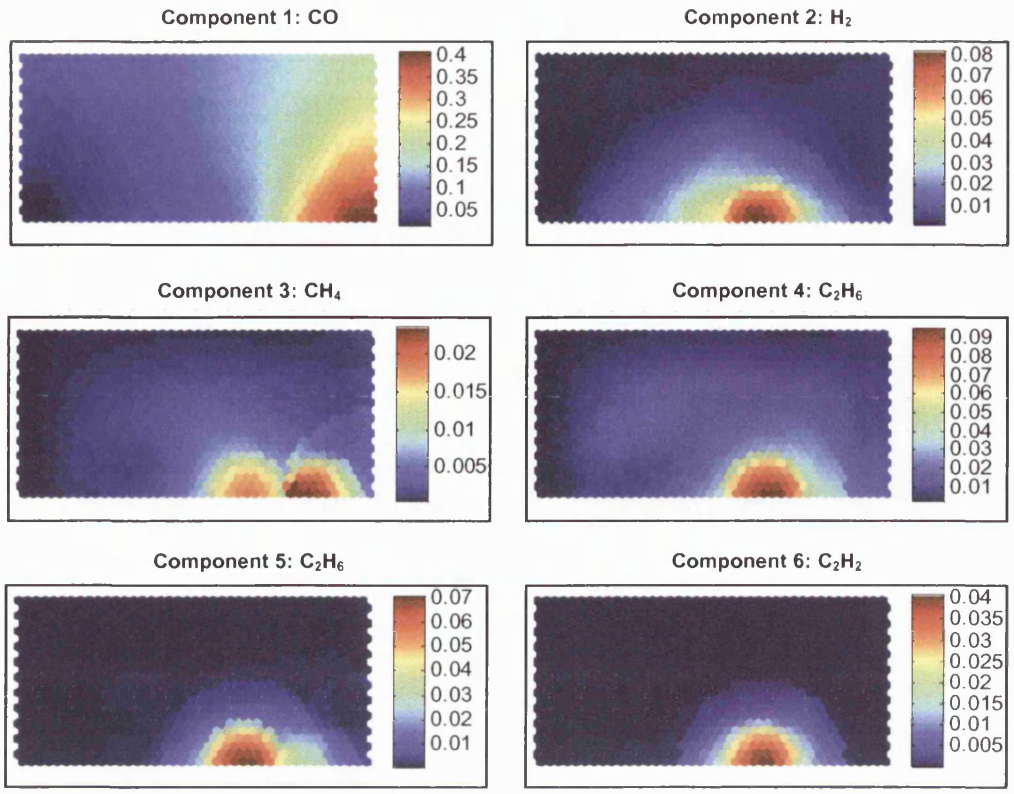


Figure 8.4: Component planes of the optimum SOM for Configuration 3 of the training data (i.e. Set C with "range" scaling)

Table 8.5: The map sizes, number of training iterations and SOM quality measures for optimum SOMs of the "bushings" DGA data

No.	Training Data	Map Size [a] [*x-dim*, *y-dim*]	Number of Training Iterations	SOM Quality Measures [d]	
				AQE [b]	TE [c]
1	Configuration 1	[40 22]	38152 (19 passes of all training vectors)	0.056	0.041
2	Configuration 2	[46 20]	40160 (20 passes of all training vectors)	0.023	0.034
3	Configuration 3	[40 22]	40160 (20 passes of all training vectors)	0.011	0.010
4	Configuration 4	[50 18]	40160 (20 passes of all training vectors)	0.006	0.025

Note:

a. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
b. AQE: Average quantisation error
c. TE: Topographic error
d. AQE and TE at the specified number of training iterations during the fine-tuning phase.

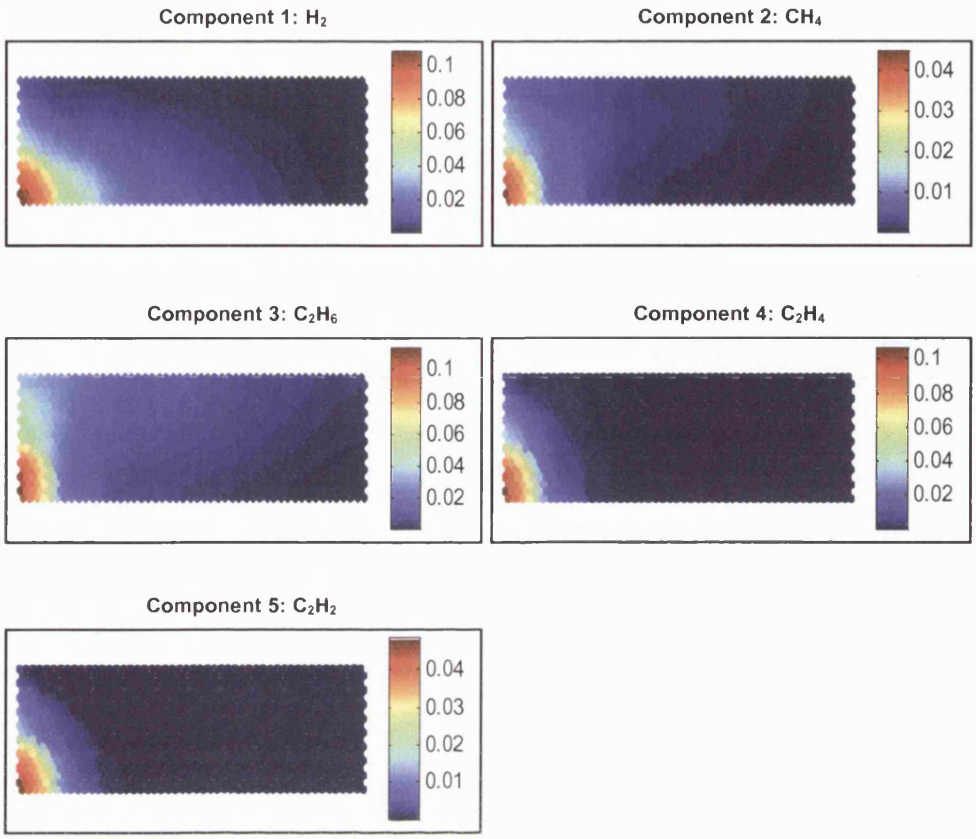


Figure 8.5: Component planes of the optimum SOM for Configuration 4 of the training data (i.e. Set D with "range" scaling)

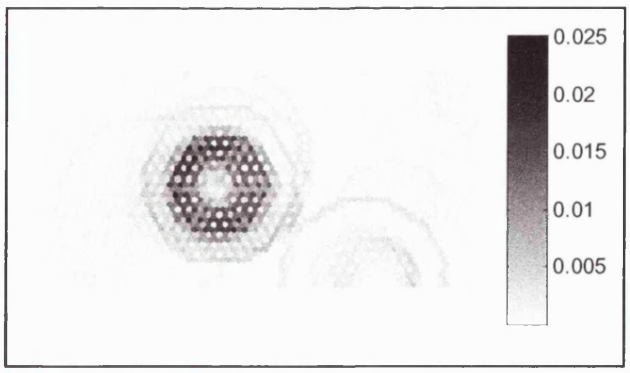


Figure 8.6: U-matrix of the optimum SOM for Configuration 1 of the training data (i.e. Set A with "range" scaling)

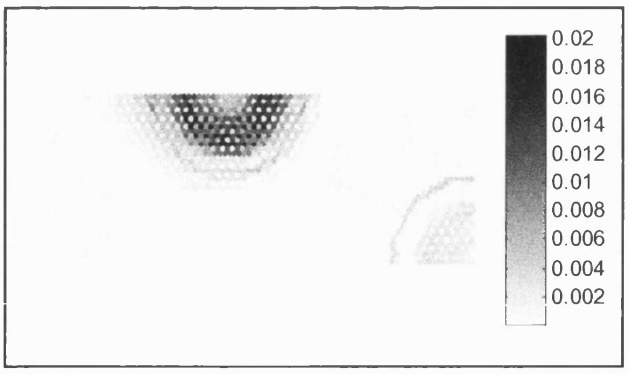


Figure 8.7: U-matrix of the optimum SOM for Configuration 2 of the training data (i.e. Set B with "range" scaling)

8.8 Analysis on Revealed Features of the "Bushings" DGA Data

As reported in the previous section, two major revealed features of the "bushings" DGA data are observed. The first revealed feature corresponds to [H_2 CH_4 C_2H_6 C_2H_4 C_2H_2] patterns and the second revealed feature corresponds to [CO CH_4] patterns. Further statistical analysis was subsequently conducted on these revealed features in order to comprehend their meaning and to possibly relate these features with the health and condition of transformer bushings.

Similar to the statistical analysis on revealed features of the DGA data of power transformers, the following steps were adopted for the analysis:

- Step 1: Identification of clusters through colour coding on the u-matrix.
- Step 2: Identification of those training vectors that regard the SOM neurons in each colour-coded cluster as their best matching units (BMUs).
- Step 3: Calculation of the average concentration (which are measured in PPM) of key dissolved gases, i.e. CO, H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, for each colour-coded cluster from training vectors that have been gathered in Step 2.

Apart from the analysis on identified clusters, the region in which no apparent cluster is formed by the key-dissolved gases was also investigated and similar procedures of analysis were performed on this region so as to determine the average concentration of key dissolved gases in this region as well.

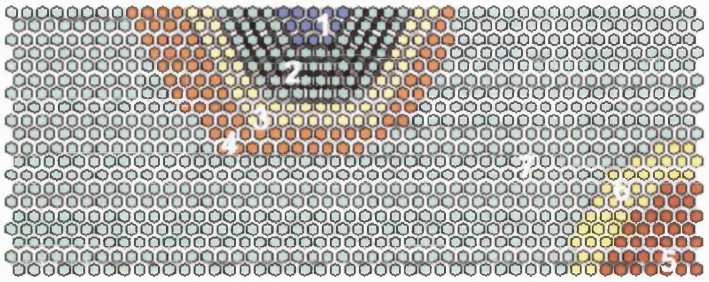
Moreover, specific attention is brought to the setting of "mask", which is the weighting factor for each component of reference vectors used during the calculation of Euclidean distances. This is due to the significantly higher scaled values of CO (which is around "0.x") when compared with the scaled values of other key dissolved gases (which are around "0.0x"). Consequently, the weighting factor for CO was set lower in order to eliminate the dominance of CO in BMU searches, as shown in Table 8.6. Note that "masks" of 1 were used for key dissolved gases of the DGA data of power transformers owing to the identical range of scaled values observed from component planes of these dissolved gases.

Table 8.6: "Masks" for key dissolved gases in BMU searches

No.	Dissolved Gas	Mask [a]
1	CO	0.01 [b]
2	H_2	1
3	CH_4	1
4	C_2H_6	1
5	C_2H_4	1
6	C_2H_2	1

Note:
a. "Masks" for non-key dissolved gases are set to 0.
b. A "mask" of 0.01 is used instead of 0.1, so as to take into account the squaring effect of each component during the calculation of Euclidean distances.

For illustration purposes, the u-matrix of Configuration 2 of the training data (i.e. Set B with "range" scaling), as illustrated in Figure 8.7, was used for performing the detailed statistical analysis. The colour-coded u-matrix and the distribution of average concentration of key dissolved gases in various colour-coded regions are illustrated in Figure 8.8.



Colour coding:

1. 1, 2, 3, 4: [H_2 CH_4 C_2H_6 C_2H_4 C_2H_2] clusters
2. 5, 6: [CO CH_4] clusters
3. 7: Apparent cluster not identified

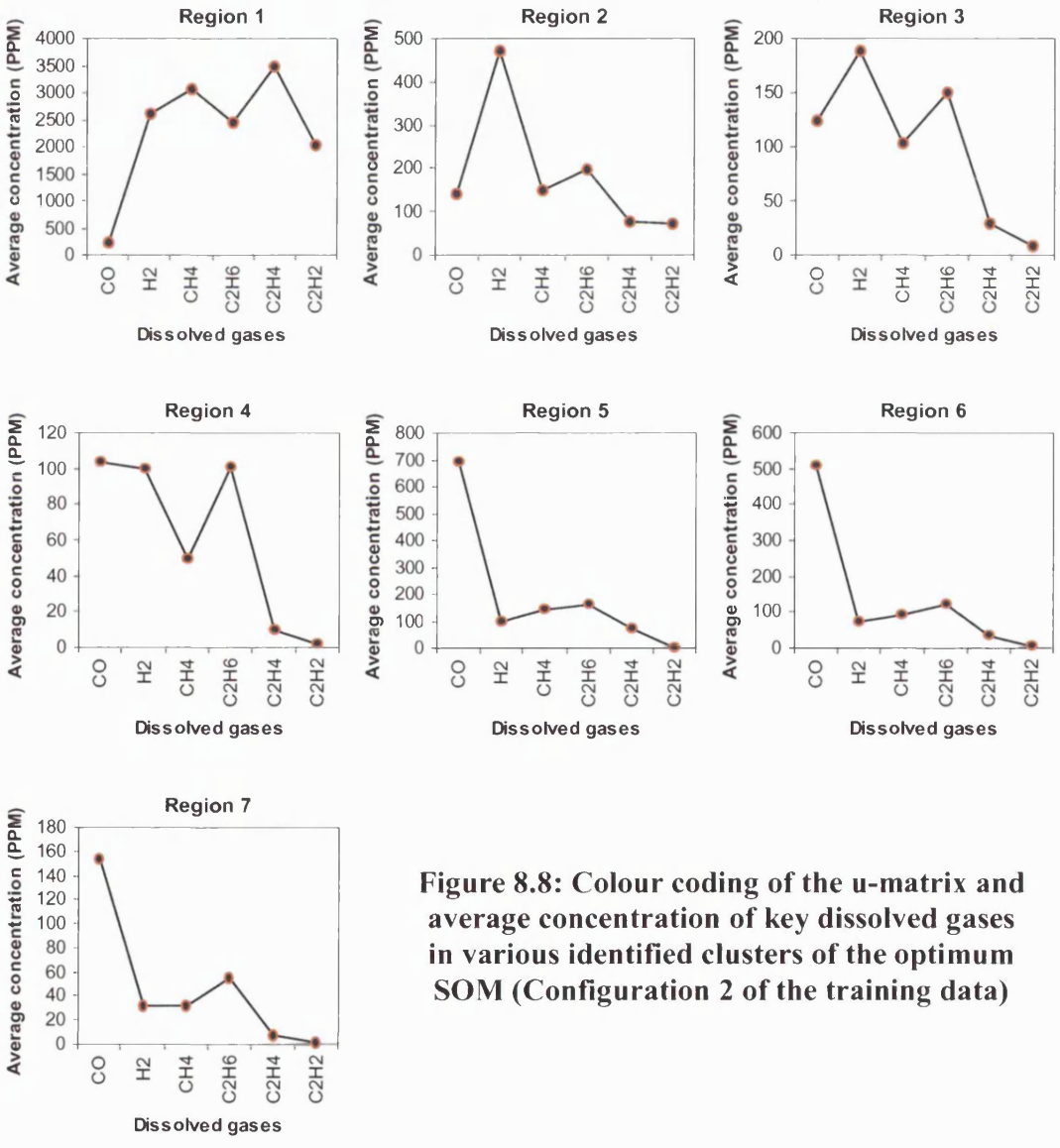


Figure 8.8: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM (Configuration 2 of the training data)

Several important characteristics can be derived from Figure 8.8:

- The average concentration of CO is low in Regions 1, 2, 3 and 4 if compared to that of Regions 5 and 6.
- The average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 increases from Regions 4 to 1.
- The average concentration of CO is very high in Regions 5 and 6.
- A consistently low average concentration of C_2H_2 is observed in Regions 5 and 6.
- The average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 is very low in Regions 7. In addition, the average concentration of CO is closer to that of Regions 1, 2, 3 and 4 but is significantly lower than that of Regions 5 and 6.

Detailed investigation was performed on those training vectors that correspond to Regions 1, 2, 3 and 4, owing to the observed dissimilarity in the distribution pattern of average concentration of key dissolved gases in these regions. It was found that those training vectors actually contain various dissimilar compositions of key dissolved gases. However, those training vectors were found to contain relatively low concentration of CO, which is the only common characteristic among those training vectors. Further SOM analysis was subsequently conducted on those training vectors in an attempt to further unearth the inherent characteristics from those data, as will be presented in the next section of the chapter.

The distribution pattern of average concentration of key dissolved gases in Regions 5 and 6 is identical to one another, in which the average concentration of CO is very high and the average concentration of C_2H_2 is consistently low when compared with the higher average concentration of H_2, CH_4, C_2H_6 and C_2H_4.

8.9 Further SOM Analysis on Regions 1, 2, 3 and 4

The SOM was further applied for the analysis of those training vectors that correspond to Regions 1, 2, 3 and 4 of Figure 8.8 for reasons described in the previous section.

The "mask" for CO was again set to 0.01 while "masks" for other key dissolved gases were set to 1 in the BMU searches during the training process of SOM. This is so because the previous observation in Regions 1, 2, 3 and 4 was attained as a result of the adoption of these "masks". In addition, scaled values of training vectors were applied directly for the training of SOM so as to ensure a consistent adoption of scaling factors from the first level (as reported in Section 8.6) to the second level (as reported in this section) of the SOM analysis.

A SOM was thus trained using those training vectors that correspond to Regions 1, 2, 3 and 4 of Figure 8.8, which have been collected via the means of best-matching. The optimum SOM was chosen according to the procedures outlined in Section 5.5.2.4 and is illustrated in Figure 8.9. Finally, the map size, number of training iterations and SOM quality measures of the optimum SOM are shown in Table 8.7 below:

Table 8.7: The map size, number of training iterations and SOM quality measures of the optimum SOM

No.	Training Data	Map Size [a] [*x-dim*, *y-dim*]	Number of Training Iterations	SOM Quality Measures [d]	
				AQE [b]	TE [c]
1	Training vectors that correspond to Regions 1, 2, 3 and 4	[32 14]	48500 (100 passes of all training vectors)	0.015	0.021

Note:
a. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
b. AQE: Average quantisation error
c. TE: Topographic error
d. AQE and TE at the specified number of training iterations during the fine-tuning phase.

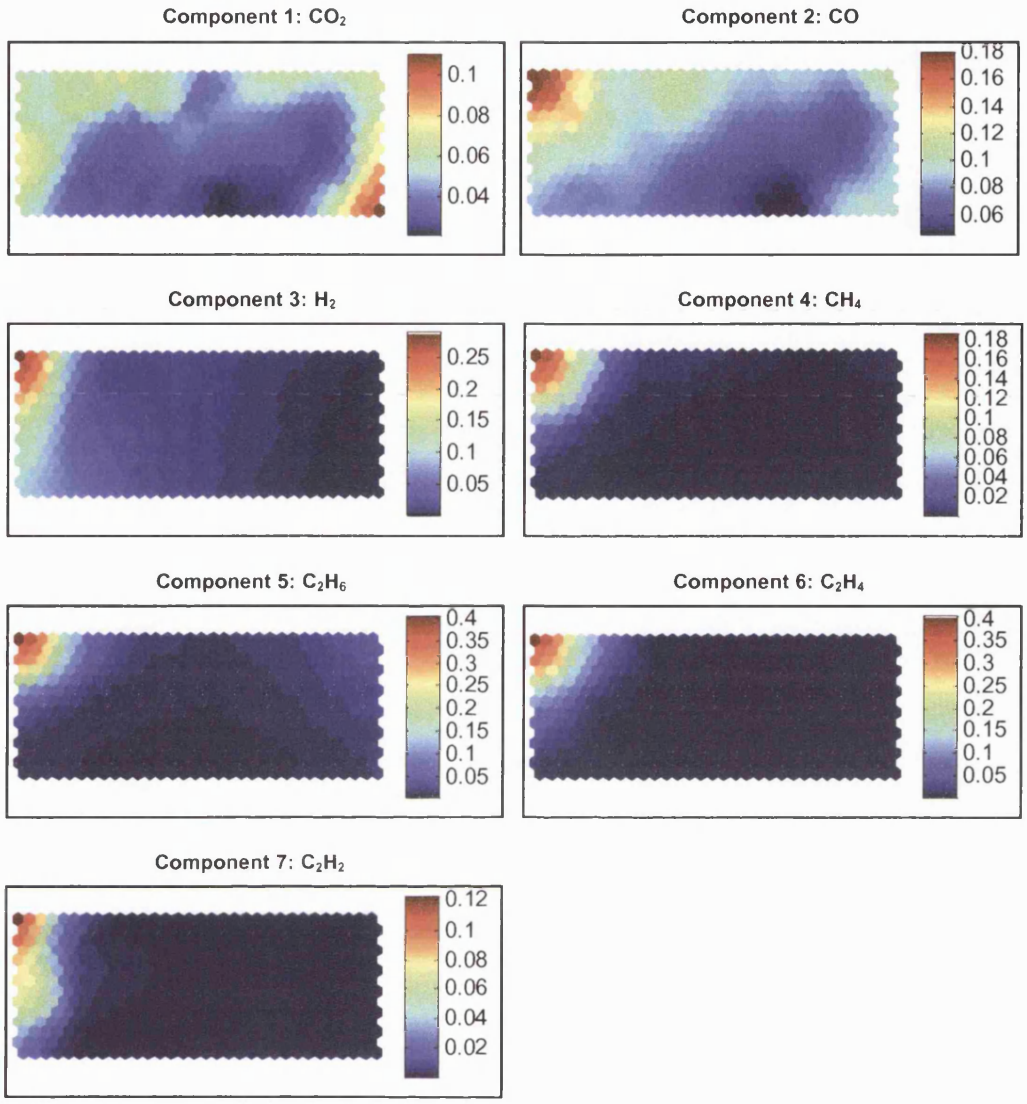


Figure 8.9: Component planes of the optimum SOM for training vectors that correspond to Regions 1, 2, 3 and 4 illustrated in Figure 8.8

As observed from Figure 8.9, the component plane of C_2H_2 actually contains two apparent patterns; the first pattern is correlated with H_2, CH_4, C_2H_6 and C_2H_4 while the second pattern corresponds to lower concentration of H_2. Furthermore, an interception is also observed between these two patterns. Finally, revealed features of those training vectors that correspond to Regions 1, 2, 3 and 4 are summarised in Figure 8.10 using the u-matrix illustration.

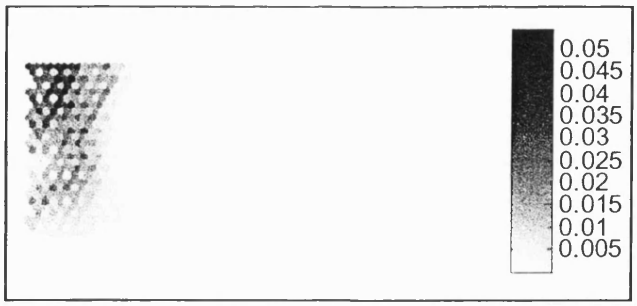
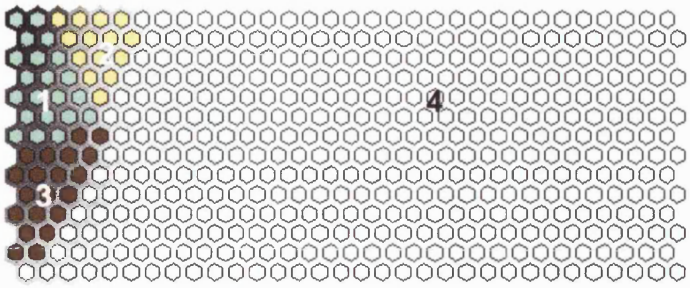


Figure 8.10: U-matrix of the optimum SOM for training vectors that correspond to Regions 1, 2, 3 and 4 illustrated in Figure 8.8

Further statistical analysis was conducted on these revealed features so as to ascertain their meaning. Hence, the u-matrix as illustrated in Figure 8.10 was partitioned and colour-coded into four regions; those training vectors that regard the SOM neurons in each colour-coded region as their BMUs were gathered and the average concentration of every key dissolved gas was then calculated. Note that the "mask" of CO was set to 0.01 while the "masks" for other key dissolved gases were set to 1 in BMU searches. The outcome of the statistical analysis is illustrated in Figure 8.11.

Several important characteristics can be derived from Figure 8.11:

- Region 1 is found to contain extremely high average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2. However, on more detail inspection on those training vectors that correspond to Region 1, it was found that these training vectors actually contain a mixture of extremely high concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, in which the distribution patterns are not as easily distinguished as patterns of Regions 2 and 3.
- CH_4, C_2H_6 and C_2H_4 are the dominant dissolved gases in Region 2. However, the average concentration of these gases is lower than that of Region 1. Therefore, the pattern of distribution of the average concentration of key dissolved gases actually resembles that of the "thermal fault" phenomenon.



Colour coding:

1. 1: Interception cluster of (2) and (3)
2. 2: [H_2 CH_4 C_2H_6 C_2H_4 C_2H_2] cluster
3. 3: [H_2 C_2H_2] cluster
4. 4: Apparent cluster not identified

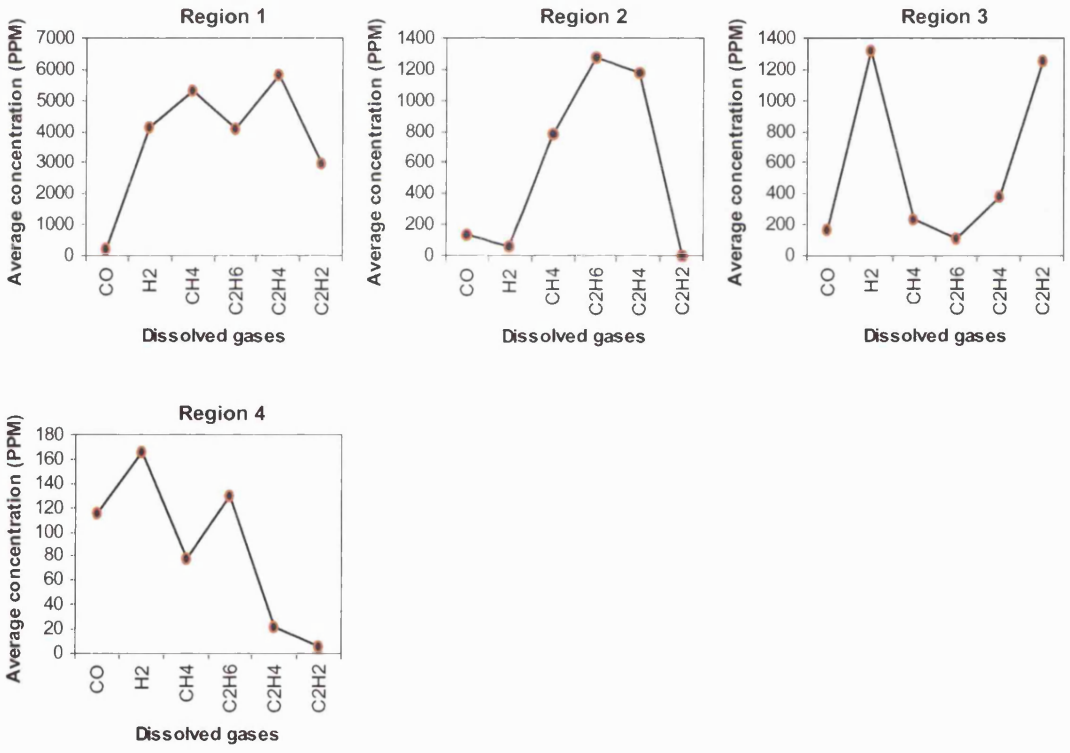


Figure 8.11: Colour coding of the u-matrix and average concentration of key dissolved gases in various identified clusters of the optimum SOM for training vectors that correspond to Regions 1, 2, 3 and 4 illustrated in Figure 8.8

- H_2 and C_2H_2 are the dominant dissolved gases in Region 2. However, the average concentration of these gases is lower than that of Region 1. Therefore, the distribution pattern of the average concentration of key dissolved gases actually resembles that of the "discharges" phenomenon.
- In Region 4, a low average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 is observed if compared to that of Regions 1, 2 and 3. However, on more detail inspection on those training vectors that correspond to this region, it is found that the majority of these training vectors typically contain higher average concentration of H_2, CH_4, C_2H_6 and C_2H_4 if compared to that of C_2H_2. However, the distribution patterns of key dissolved gases in this region are not as easily distinguishable as patterns of Regions 2 and 3.
- All regions are found to contain a low average concentration of CO, which actually matches the previous observation in Section 8.8.

8.10 Hypothetical Association of Revealed Features

Although the revealed features of the "bushings" DGA data are not as apparent and comprehensible as that of the DGA data of power transformers, hypothetical association can still be established with several conditions of transformer bushings, albeit in a less confident manner when compared with revealed features of the DGA data of power transformers.

As reported in Section 8.8, the first level of SOM analysis has revealed two major features. The first feature resembles that of the "thermal fault" (TF) phenomenon with involvement of cellulose insulation, due to the high average concentration of CO and high average concentration of CH_4, C_2H_6 and C_2H_4 if compared to that of C_2H_2. On the contrary, the second feature corresponds to very low average concentration of CO and a mixture of distribution patterns of other key dissolved gases. Besides, a region of low average concentration of all key dissolved gases was also discovered, which can be related to the normal operating condition of transformer bushings.

As reported in Section 8.9, second level of SOM analysis was subsequently performed on those training vectors that correspond to the second feature as mentioned above. Consequently, four further features were unearthed. The first feature resembles that of severe "discharges" and/or severe TF phenomena, due to the extremely high average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 or a mixture of these gases. In addition, the second feature resembles that of the severe TF phenomenon due to the dominance of CH_4, C_2H_6 and C_2H_4 while the third feature resembles that of the severe "discharges" phenomenon due to the dominance of H_2 and C_2H_2. Finally, the fourth feature resembles that of less severe TF and/or less severe "electrical fault" (i.e. EF, which includes discharges and partial discharges (PD)) phenomena, owing to the considerably lower average concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 or a mixture of these gases. It is, however, certain that there is no involvement of cellulose insulation in the abovementioned faults owing to the consistently low average concentration of CO.

The hypothetical association of various revealed features is illustrated in Figure 8.12.

8.11 Validation of the Hypothesis

As discussed in Section 8.3, it is less critical to identify the "type" of a suspected fault in transformer bushings and monitor its evolution owing to the significantly lower cost of replacing a failed bushing compared to the cost of repairing or replacing a failed main unit of power transformer. In fact, engineers are far more interested in the detection of the "presence" of faults rather than in the identification of the "type" of a suspected fault.

Consequently, there is only one known approach, described herein, and as outlined in the IEC 60599 Standard [20], for relating the DGA data of bushings (i.e. not specifically transformer bushings) with several common incipient faults of bushings.

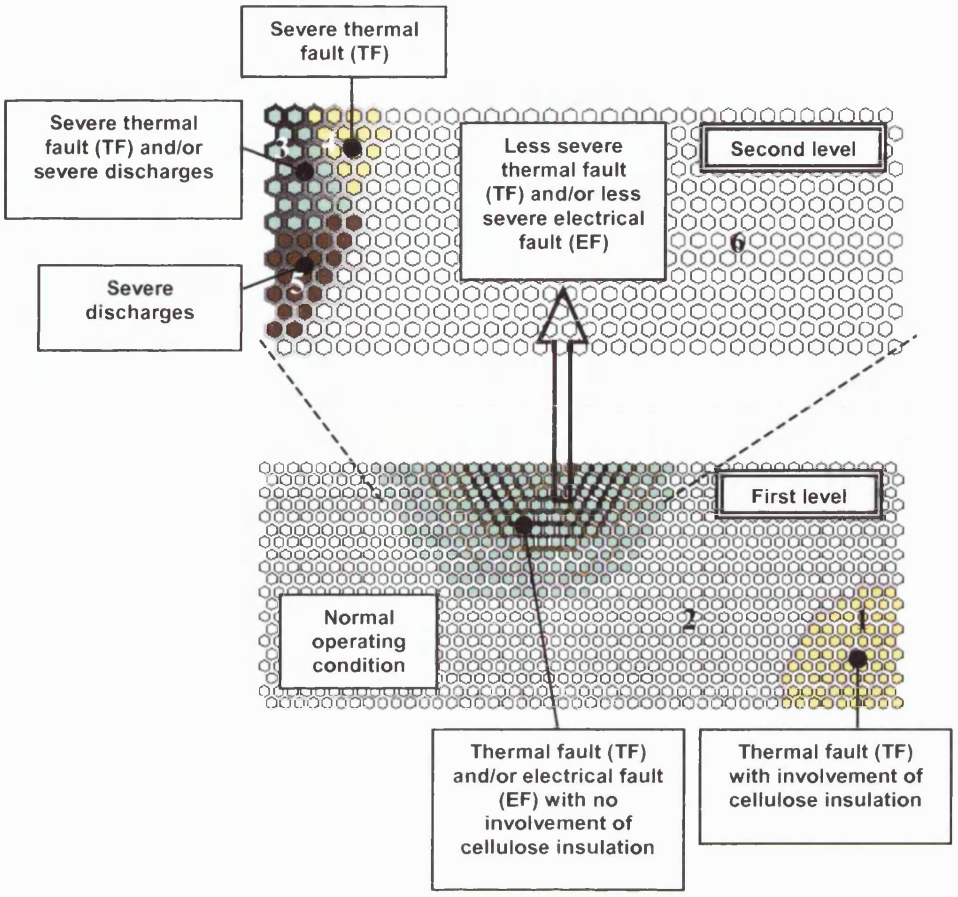


Figure 8.12: Association of revealed features with conditions of transformer bushings

In addition, several fault cases are also available, which are the DGA data of transformer bushings that have actually failed and causes of such failures have been identified through post-mortem examinations. However, only one DGA data is available for each case as distinct from the entire DGA history of power transformers.

The hypothesis presented in Section 8.10 was thus compared with the IEC's method and with actual findings of NGC's engineers, as will be presented in following sections.

8.11.1 Comparison with IEC's Approach

IEC's method for interpretation of the DGA data of bushings is very similar to the interpretation for the DGA data of power transformers, which has been presented in Section 3.4.3. However, for the case of bushings, the value of CH_4/H_2 ratio is assigned as 0.07 rather than 0.1 as for the case of power transformers for the detection of partial discharges (PD). The method of interpretation for bushings is summarised in Table 8.8.

Table 8.8: IEC's method for interpretation of the DGA data of bushings

Case	Characteristic Fault	$\frac{C_2H_2}{C_2H_4}$	$\frac{CH_4}{H_2}$	$\frac{C_2H_4}{C_2H_6}$
PD	Partial discharges	NS [b]	< 0.07	< 0.2
D1	Discharges of low energy	> 1	0.1 – 0.5	> 1
D2	Discharges of high energy	0.6 – 2.5	0.1 – 1	> 2
T1	Thermal fault (T < 300 °C)	NS [b]	> 1 but NS [b]	< 1
T2	Thermal fault (300 °C < T < 700 °C)	< 0.1	> 1	1 – 4
T3	Thermal fault (T > 700 °C)	< 0.2 [c]	> 1	> 4

Note:
a. The above ratios are significant and should be calculated only if at least one of the gases is at a concentration and at a rate of gas increase above typical values.
b. Not significant whatever the value.
c. An increasing value of C_2H_2 may indicate that the hotspot temperature is higher than 1000 °C.

This method was implemented in a manner similar to that for power transformers, as has been presented in Section 4.6. In addition, the "normality" test for the DGA data of transformer bushings is based on the typical concentrations adopted by engineers at the NGC, as shown in Table 8.9.

Table 8.9: Typical concentrations for dissolved gases of bushings

Dissolved Gas	Typical Concentration (PPM)
H_2	140
CH_4	40
C_2H_6	70
C_2H_4	30
C_2H_2	0.3

Therefore, those training vectors that correspond to each numbered region as shown in Figure 8.12 were subjected to interpretation by the IEC's method, of which the results are presented in Table 8.10. Note that only a general distinction of faults is considered since the severity of these faults can be observed easily from those training vectors that correspond to each region.

Table 8.10: Comparison of hypothesis with IEC's method

		Interpretation by IEC's method				
Region [Number of Training Vectors]	**Hypothesis**	**Normal**	**EF**		**TF**	**N/I [a]**
			D	**PD**		
1 [54]	Thermal fault (TF) with involvement of cellulose insulation	1			53	
2 [1469]	Normal	771	3	1	681 [b]	13
3 [8]	Severe thermal fault (TF) and/or severe discharges (D)		3		3	2
4 [3]	Severe thermal fault (TF)				3	
5 [8]	Severe discharges (D)		7	1		
6 [466]	Less severe thermal fault (TF) and/or less severe electrical fault (EF)	19	8	20	408	12

Note:

a. N/I: No interpretation.

b. The majority of DGA data that is interpreted as TF in the hypothetical "normal" region is actually belong to the lowest TF category, i.e. case T1 of Table 8.8.

The comparison table as illustrated in Table 8.10 can be analysed in the following manner:

- IEC's method interprets 53 out of 54 DGA data in Region 1 as "TF", and this matches very well with the hypothesis. Although not proven using IEC's method, the involvement of cellulose insulation is clear owing to the considerably higher concentration of CO in Region 1.
- IEC's method interprets 771 out of 1469 DGA data in Region 2 as "normal" and 681 out of 1469 DGA data as "TF". However, those DGA data that are interpreted as "TF" actually correspond to the lowest TF category, i.e. case T1 in Table 8.8. In addition, it was later informed by engineers of the NGC that typical concentrations as shown in Table 8.9 were actually taken from IEC's recommendation with amendment to the typical value of C_2H_2. It is thus possible that the use of different typical concentrations may alter the number of DGA data that is interpreted as "normal" in Region 2.
- IEC's method interprets 6 out of 8 DGA data in Region 3 as either "discharges" or TF, and this matches well with hypothesis. In addition, IEC's method cannot provide interpretation for two DGA data in this region. Although not proven using IEC's method, the high severity of these faults is certain owing to the extremely high concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 or a mixture of these gases in this region.
- Interpretations provided by the IEC's method for Regions 4 and 5 matches very well with the hypothesis. In addition, the high severity of faults is certain owing to the very high concentration of respective dominant dissolved gases in these regions.
- IEC's method interpreted 408 out of 466 DGA data in Region 6 as "TF", which again matches the hypothesis. Although not proven using IEC's method, it is certain that the TF phenomenon is of a less severe characteristic, owing to the considerably lower concentration of CH_4, C_2H_6 and C_2H_4 if compared to that of Regions 3 and 4.

Therefore, the suggested hypothesis in Section 8.10 compares quite well with IEC's method. Although only a general comparison is performed, the severity of identified faults is easily proven through direct inspection on the training vectors.

8.11.2 Validation by Using Actual Fault Cases

Since a transformer bushing is normally replaced once the presence of a fault is suspected based on the concentration of certain dissolved gases, the availability of actual fault cases is thus very rare. The two fault cases presented herein are real DGA data of transformer bushings that have actually failed and their causes of failures have been confirmed through post-mortem examinations. However, only one DGA data is available for each case, as shown in Table 8.11, which are the concentration of nine dissolved gases and moisture after the failure had occurred.

Table 8.11: The concentration of dissolved gases for two actual fault cases

Concentration (PPM)	Case 1	Case 2
H_2O	4	4
N_2	53675	55566
O_2	5502	2105
CO_2	623	364
CO	349	520
H_2	9678	9354
CH_4	376	1225
C_2H_6	805	1177
C_2H_4	978	1177
C_2H_2	3322.7	4906.8

These two fault cases were "mapped" onto two levels of u-matrix as illustrated in Figures 8.7 and 8.10 via the means of best-matching. The proposed health and condition of each transformer bushing, as discerned from the "location" of the BMU of each DGA data on the u-matrices, is then compared with the actual findings of engineers at NGC.

Firstly, the location of the BMU of Case 1 on two levels of u-matrix is illustrated in Figure 8.13. As can be seen from the figure, the BMU is located within the "severe

TF and/or severe discharges" region. This matches well with the actual cause of failure, which is discharges between the outer foil of bushing core and the flange as a result of a floating potential, and this was a direct consequence of a loose screw causing the rotation of bushing core that subsequently snapped the test-tap wire connection.

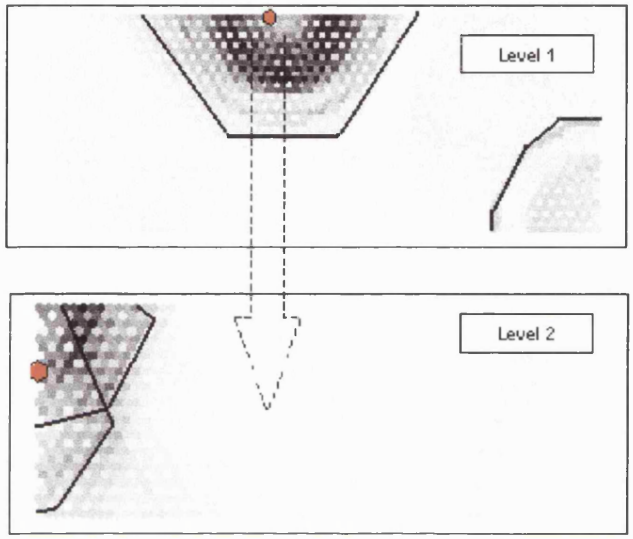


Figure 8.13: The location of BMU for Case 1 on two levels of u-matrix

Secondly, the location of the BMU for Case 2 on two levels of u-matrix is illustrated in Figure 8.14. As seen from the figure, the location of BMU is identical, which is on the "severe TF and/or severe discharges" region. Again, this matches well with the actual cause of failure, which is arcing in oil at the high voltage connection lead near the top of the bushing.

The hypothesis is thus capable of providing convincing diagnosis of a fault based on the location of the BMU of a DGA data onto two levels of u-matrix. It is able to provide an easy means of fault diagnosis once a post-mortem examination is carried out. Unfortunately, not enough fault cases are available for further validation of the hypothesis.

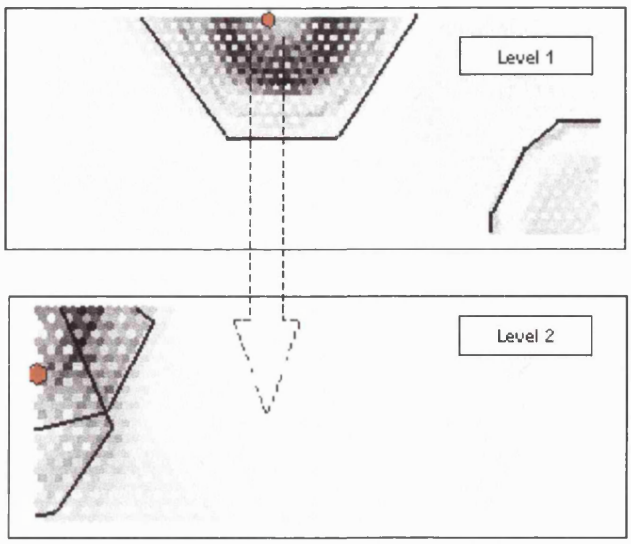


Figure 8.14: The location of BMU for Case 2 on two levels of u-matrix

8.12 Advantages of the Proposed Approach

As mentioned in Section 8.3, engineers are more interested in detecting the presence of problems or faults in transformer bushings rather than diagnosing the type of suspected faults. Nevertheless, the hypothesis as presented in Section 8.10 is still useful and beneficial to engineers owing to the following reasons:

- The proposed approach is able to provide an easy means of fault diagnosis based on the location of the BMU of DGA data of a transformer bushing on the two levels of u-matrix illustrated in Figure 8.12.
- In a situation in which the diagnosis of faults is not needed, the proposed approach can be used solely for detecting the "presence" of faults, which can be accomplished if the BMU of the DGA data of a transformer bushing falls into any one of the identified fault regions in the first level of u-matrix illustrated in Figure 8.12.
- The interpretation as provided by the proposed approach is more convincing than the IEC's method because it is more specialised and fundamentally

based on the inherent characteristics of the DGA data of transformer bushings. The IEC's method is a more generic approach applicable to all types of bushings.

- If the interval between subsequent DGA tests can be increased, engineers can monitor the evolution of a fault if detected through plotting of the DGA history of a transformer bushing onto the two levels of u-matrix illustrated in Figure 8.12.

8.13 Summary

The proposed approach, which is based on the application of SOM, has been applied for the analysis and interpretation of the DGA data of transformer bushings. Essentially, revealed features of the DGA data of transformer bushing are found to be very different from that of power transformers. However, these revealed features can be hypothetically associated with several conditions of transformer bushings, which have been validated using IEC's method and two actual fault cases. Although it is far less critical for determining the "type" of faults in real-life, the proposed approach is still useful for engineers who would like to get an early means of fault detection or fault diagnosis before a suspected faulty bushing is replaced or a post-mortem examination is performed.

CHAPTER 9

DATA MINING ON THE MULTIPLE SENSOR DATA OF A POWER TRANSFORMER USING THE SOM

9.1 Introduction

The data mining (DM) technique of self-organising map (SOM) was applied for the analysis of a set of on-line monitoring data comprising of various measurements of sensor variables collected from a unit of power transformer. The main objective of performing this study is to investigate the performance of SOM for unearthing the inherent characteristics of a set of heterogeneous condition assessment data, which contains a diverse mix of various on-line monitoring data of a power transformer.

9.2 Multiple Sensor Data of Power Transformer

The National Grid Company (NGC), UK, has been carrying out an hourly on-line monitoring process on a unit of power transformer (with voltage level of 400/132 kV and power rating of 240 MVA) for almost a year. A total of eleven variables are monitored hourly using various on-line sensors; these variables and their units are listed in Table 9.1. Consequently, 2447 hourly sensor measurements were acquired from the date of 31/8/00 to 20/8/01. These measurements are recorded in a format as illustrated in Table 9.2.

Table 9.1: Sensor variables recorded from a power transformer

No.	Sensor Variable	Abbreviation	Unit
1	Temperature – Ambient	TA	°C
2	Relative Humidity – Ambient	RHA	%
3	Temperature – Top Oil	TTO	°C
4	Temperature – Bottom Oil	TBO	°C
5	Temperature – Load Tap Changer	TLTC	°C
6	Oil Relative Saturation	ORS	%
7	Oil Temp. at Moisture Sensor	OTMS	°C
8	Phase Current	PC	A
9	Hydrogen Content	H_2	PPM
10	Cooling Circuit 1 – Fans	FAN	ON (1)/OFF (0)
11	Cooling Circuit 2 – Pumps	PUMP	ON (1)/OFF (0)

Table 9.2: Format of the multiple sensor data

A	B	C	D	E	F
Recording Time Stamp (Date – Time)	**TA (°C)**	**RHA**	**TTO (°C)**	**TBO (°C)**	**TLTC (°C)**
07/12/2001 09:02	16.9778	61.2312	38.9	39.8	32.4
07/12/2001 10:02	18.5228	57.1304	39.6	40.4	32.6
07/12/2001 11:02	16.377	65.6181	40.2	40.9	33

G	H	I	J	K	L
ORS	**OTMS (°C)**	**PC (A)**	**H_2 (PPM)**	**FANS (ON/OFF: 0/1)**	**PUMPS (ON/OFF: 0/1)**
10.975	30.455	391	148.4	0	1
10.93	30.928	402	148	0	1
10.96	31.148	409	148	0	1

However, the acquired data is not complete due to the previous failure of the data-logging equipment. Therefore, some portions of the sensor data are missing and various "blanks" (i.e. no data entries or records) are observed within the data. The breakdown of the acquired sensor measurements according to the date and amount of data is shown in Table 9.3. Despite the foregoing imperfections of the data, it is not possible to collect more measurements from other power transformers since NGC has only carried out the on-line monitoring practice on one unit of a power transformer. Hence, results from subsequent analysis should be treated with caution due to the foregoing imperfections of the data.

Table 9.3: Breakdown of acquired sensor measurements according to date and amount of data

No.	Date	Amount of Data
1	31/9/00 to 19/10/00	1171
2	17/1/01 to 24/1/01	168
3	22/3/01 to 25/3/01	73
4	28/3/01	1
5	24/5/01 to 27/5/01	74
6	2/7/01 to 20/08/01	960

9.3 Pre-Processing of the Multiple Sensor Data

As mentioned previously, a lot of "blanks" are present within the sensor data due to the previous failure of the data-logging equipment. Instead of removing the entire line of measurements of sensor variables of which one of the measurements is "blank", the "blank" was replaced by the notation "NaN" (i.e. Not-a-Number). The SOM is capable of handling the "NaN" by ignoring those components of training vectors that contain "NaNs" for the calculation of Euclidean distances during the SOM training process. Hence, the original information contained within the multiple sensor data can be retained as much as possible since measurements of other sensor variables can still be used for the training of SOM.

9.4 Data Mining on the Multiple Sensor Data

A total of eleven input components are available for the training of SOM, which are from various sensor variables as illustrated in Table 9.1. In addition, the training data was transformed using the "range" method due to its proven usefulness for data analysis; the "range" scaling method is described in the Appendix of the thesis.

In addition, configuration and training parameters of SOM as shown in Table 7.3 were adopted during the training process, based on the fact that the SOM technique is generic and identical configuration and training parameters can be utilised for any set of data. Finally, the optimum SOM was selected out of a number of trained SOMs according to procedures as described in Section 5.5.2.4.

9.5 Optimum SOM of the Multiple Sensor Data

The optimum SOM of the multiple sensor data is visualised using the component-plane illustration, as shown in Figure 9.1. In addition, the map size, number of training iterations and SOM quality measures for this optimum SOM are illustrated in Table 9.4 below.

Table 9.4: The map size, number of training iterations and SOM quality measures of the optimum SOM

No.	Training Data	Map Size [a] [*x-dim*, *y-dim*]	Number of Training Iterations	SOM Quality Measures [d]	
				AQE [b]	TE [c]
1	Multiple sensor data	[46 22]	61175 (25 passes of all training vectors)	0.1599	0.0826

a. *x-dim*, *y-dim*: Number of neurons in the *x*- and *y*-dimension of SOM.
b. AQE: Average quantisation error
c. TE: Topographic error
d. AQE and TE at the specified number of training iterations during the fine-tuning phase.

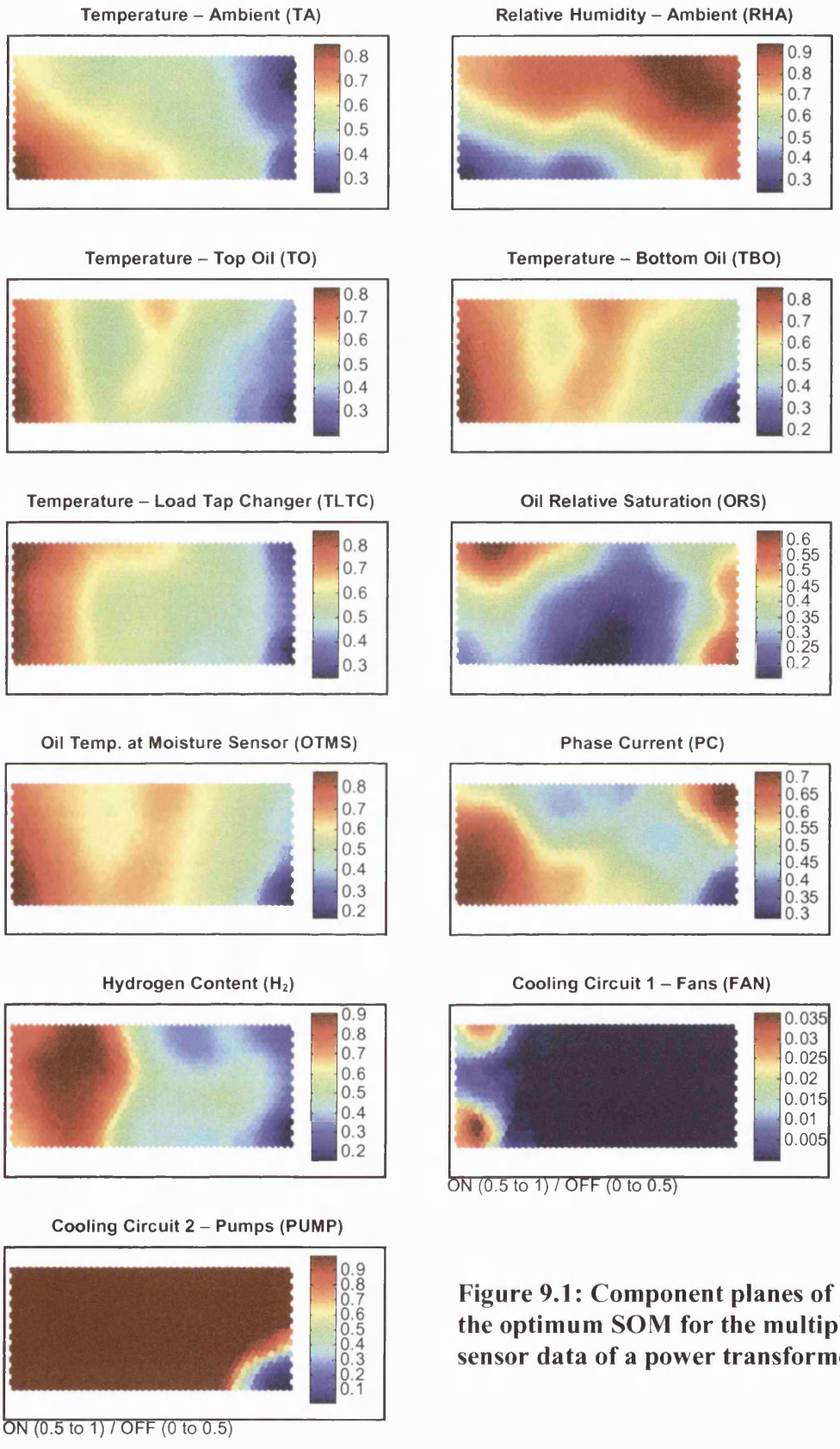


Figure 9.1: Component planes of the optimum SOM for the multiple sensor data of a power transformer

Several conclusions can be drawn from Figure 9.1:

- All sensor variables that are based on the measurement of temperature, i.e. TA, TTO, TBO, TLTC and OTMS, are observed to have quite similar characteristics, judging from the fact that the magnitudes of revealed patterns for these variables actually increase from right to left.
- The ambient temperature (i.e. TA) is regarded as one of the key influential factors for the variation of other temperature variables.
- The revealed pattern of RHA is in direct opposition to that of TA, which is understandable since the ambient humidity is normally lower in the case of higher ambient temperature.
- The relationship of ORS with other sensor variables is unclear.
- Three regions of interest can be observed from the PC component-plane. The correspondence of other sensor variables with these regions of PC is found to be both interesting and comprehensible:
  a. High-magnitude of PC in the lower-left part of the component plane corresponds to high-magnitude of TA, TTO, TBO, TLTC and OTMS. In addition, RHA is of low-magnitude in that region and the concentration of H_2 is high. Finally, the FAN is OFF and the PUMP in ON.
  b. High-magnitude of PC in the upper-right part of the component plane corresponds to low-magnitude of TA, TTO, TBO, TLTC and OTMS. In addition, RHA is of high-magnitude in that region and the concentration of H_2 is low. Finally, the FAN is OFF and the PUMP is ON.
  c. Low-magnitude of PC in the lower-right part of the component plane corresponds to low-magnitude of TA and TLTC and the lowest-magnitude of TTO, TBO and OTMS. In addition, RHA is of high-magnitude in that region and the concentration of H_2 is low. Finally, the FAN is OFF and the PUMP is OFF.
- The foregoing observations on three regions of the PC component-plane can be summarised based on Figure 9.2.

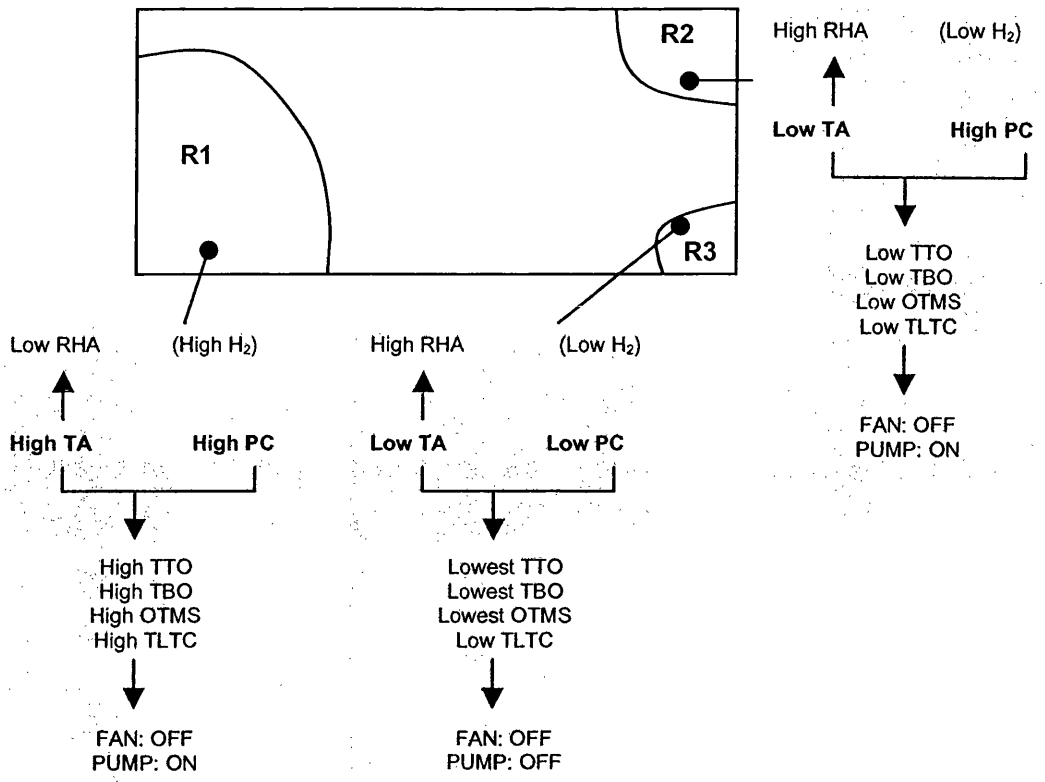


Figure 9.2: Correspondence of various sensor variables with three regions of interest of the PC component-plane

As can be deduced from the foregoing findings on revealed patterns of the multiple sensor data, two key influential factors for the variation in magnitudes of other sensor variables are the ambient temperature (i.e. TA) and the phase current of power transformer (i.e. PC). Since the phase current is directly proportional to the loading of power transformer, its hourly variation couples with the hourly variation of ambient temperature could impact on the variation in magnitudes of other sensor variables.

The SOM has thus successfully unearthed the inherent characteristics of the multiple sensor data and presented these revealed features in a discernible and comprehensible format, as has been illustrated in Figure 9.1. In fact, these revealed features of the multiple sensor data can actually be applied for monitoring the hourly operating condition of power transformer, as will be explained in the next section.

9.6 On-Line Monitoring of Power Transformer based on the Multiple Sensor Data

Since the hourly operating condition of the power transformer is implicitly signified by the variation in magnitudes of various sensor variables, it can hence be monitored via the means of visualising the movement of a trajectory on various component planes of sensor variables.

To form the trajectory, the best-matching-unit (BMU) of all sensor measurements at a particular time is first determined, based on the minimum Euclidean distance of all sensor measurements at a particular time with reference vectors of the optimum SOM; all BMUs that correspond to the time of 00:00 (i.e. 12 am) to 23:00 (i.e. 11 pm) are subsequently linked together to form the trajectory for a particular day. It should be noted that the location of this trajectory on various component planes of sensor variables is exactly identical since each BMU is determined based on the combined effect of all sensor measurements during the calculation of Euclidean distance.

An example of such application is demonstrated in Figure 9.3, which illustrates the movement of the trajectory on several chosen component planes of the optimum SOM for the day of 25/5/01 (time varies from 00:00 (i.e. 12 am) to 23:00 (i.e. 11 pm)). As can be observed from the figure, the hourly variation in magnitudes of various sensor variables can be easily scrutinised from these component planes. Furthermore, the relative location of the trajectory with respect to the highest and lowest regions of each sensor variable can also be identified easily.

Hence, more information on the hourly operating condition of the power transformer can be obtained based on the foregoing means of on-line monitoring when compared to the conventional means of looking at several parallel time-sequences of sensor variables. This is because magnitudes of various sensor variables for a particular day can be easily compared against the maximum and minimum recorded values.

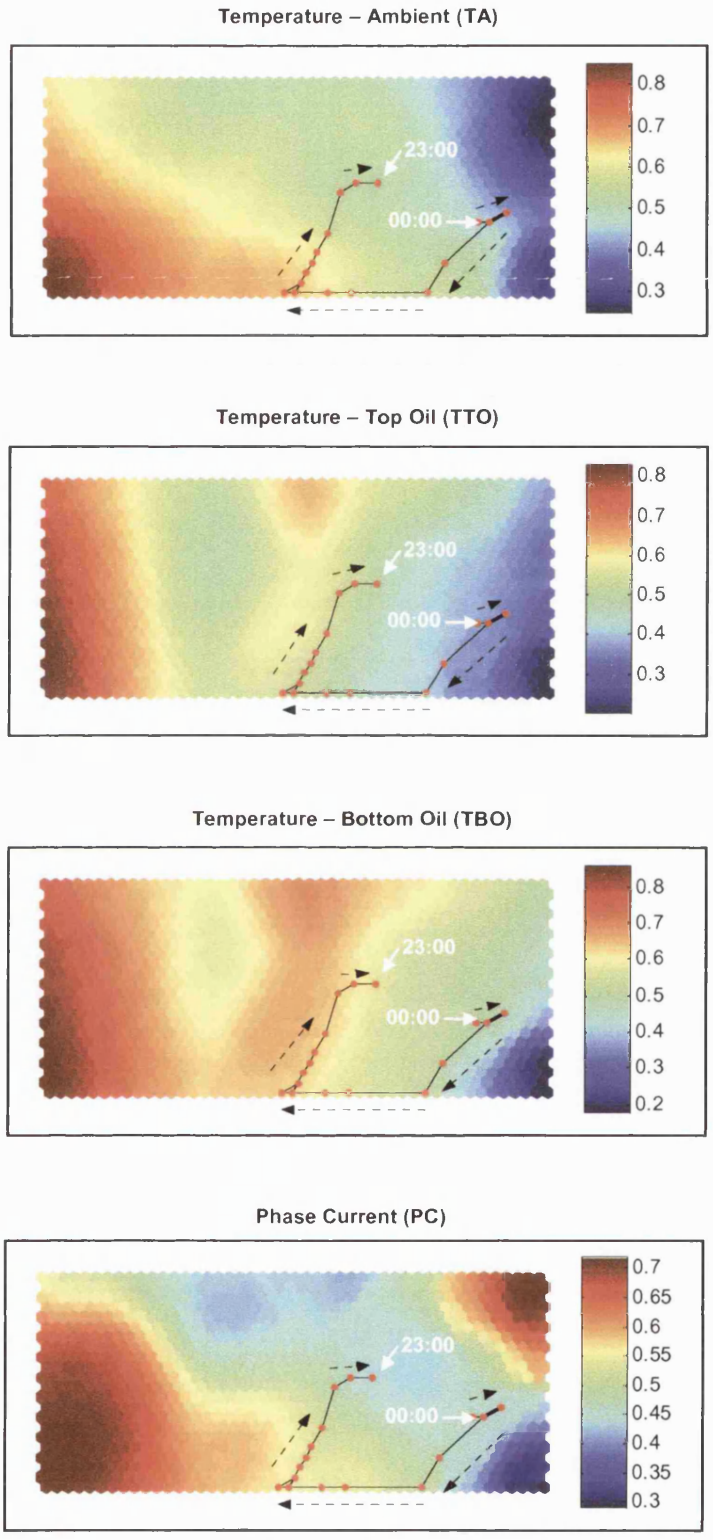


Figure 9.3: Hourly trajectory of sensor variables on the day of 25/5/01

Furthermore, the foregoing means of on-line monitoring can also be utilised for the detection of the onset of problems or faults within the power transformer. For example, if the trajectory of a power transformer condition during an extremely cold day (hence very low TA) moves into the region of which high TA is observed, some problems or faults might have occurred within the power transformer that causes the changes in magnitudes of other sensor variables which subsequently lead to the "abnormal" movement of trajectory. Unfortunately (or fortunately), this particular unit of power transformer of which the multiple sensor data was collected has been operating faultlessly since the on-line monitoring facility was installed. Hence, no actual fault cases are available to investigate the foregoing claim.

9.7 Summary

The SOM has successfully unearthed the inherent characteristics of the multiple sensor data and presented them in a discernible and understandable format. More importantly, these revealed features can be utilised for performing hourly on-line monitoring of the operating condition of a power transformer, based on the movement of the trajectory on various component planes of sensor variables. The foregoing means of on-line monitoring is more advantageous and informative when compared to the conventional means of looking at several parallel time-sequences of sensor variables; this is so since the trajectory of sensor variables on a particular day can be easily compared against the maximum and minimum recorded values and any unexpected movement of the trajectory can be regarded as a warning signal for the onset of problems or faults within the power transformer.

CHAPTER 10

CONCLUSIONS AND FUTURE WORK

10.1 Introduction

Power transformers are among the most expensive and critical components of the electrical power system. Consequently, it is important to maintain the reliability of power transformers while at the same time maximising their utilisation. These requirements are of paramount importance to power utilities, owing to the mounting challenges faced by the power utilities to supply uninterrupted power and efficient services to customers in today's increasingly competitive and deregulating energy market. Owing to the crucial role of power transformers, various condition assessment approaches to essentially identify any potential problems have been proposed. However, although a reasonably accurate recording of the condition assessment data can be obtained, there are still weaknesses when it comes to providing accurate and convincing interpretation of the condition assessment data.

In view of the foregoing, the work presented in this thesis has been devoted to the investigation and development of a novel and improved approach for the analysis and interpretation of the condition assessment data. More specifically, the major focus of the work presented herein has been placed on developing an improved approach for the analysis and interpretation of the dissolved gas analysis (DGA) data of power transformers and their bushings. In addition, the proposed approach is also applied on a limited scale for the analysis of a set of on-line monitoring data

comprising of various sensor measurements collected from a unit of power transformer.

This chapter briefly reviews all major studies that have been performed in this research project and a number of significant achievements that have been attained. Finally, possible future work with regard to this research is also proposed at the end of this chapter.

10.2 Review of Major Studies and Achievements Attained in This Research Project

Throughout the three years of research on this project, various major studies have been performed, as reported in this thesis. More importantly, several important achievements have been attained as a result. The following sections will provide a brief review of these major studies and the achievements that have been attained.

10.2.1 Implementation and Comparison of Conventional DGA Interpretation Schemes

Several established conventional interpretation schemes for the DGA data of power transformers have been implemented in this research project. Based on these implemented schemes, a meaningful comparison has been conducted in order to assess the strengths and weaknesses of these conventional schemes. It has been subsequently demonstrated that the application of different DGA schemes on an identical set of DGA data gives rise to the different interpretation on the health and condition of power transformers. In addition, there are also cases in which these conventional schemes are unable to provide any interpretation. Thus, the foregoing investigation clearly illustrates the inherent weaknesses of conventional schemes in analysing and interpreting the DGA data of power transformers.

10.2.2 Addressing Issues for the Practical Application of SOM

Although various theoretical aspects of self-organising map (SOM) have been well-defined, further issues need to be considered for its application to solve some practical problems. These issues are, for example, the setting of SOM configuration and its training parameters and procedures for the selection of optimum SOM. Hence, detailed investigations and experiments have been conducted so as to produce a systematic and effective means for tackling the foregoing issues; this has culminated in the establishment of a set of procedures for the setting of configuration and training parameters of SOM and the proper selection of optimum SOM. It has been subsequently demonstrated that these established procedures are very robust and generic, and can be applied to any set of condition assessment data very effectively.

10.2.3 Statistical Analyses of the DGA Data of Power Transformers

Various statistical analyses have been conducted on the DGA data of power transformers. These were performed so as to gain a rudimentary insight into the characteristics of the data. Specifically, the DGA data has been examined in terms of its partition according to voltage levels, power ratings and the total age of the corresponding power transformers from which oil samples were extracted. More importantly, statistical distributions of various dissolved gases have been investigated. It has been subsequently discovered that the majority of dissolved gases is normally concentrated around regions of low concentration; there are only a few rare cases where the concentration of dissolved gases is very high. The foregoing discoveries are very important for the subsequent collective analysis of the DGA data of power transformers, which forms the most important part of this research project.

10.2.4 Data Mining on the DGA Data of Power Transformers

The proposed approach, which is based on a data mining (DM) technique known as the SOM, has been applied for the analysis and interpretation of the DGA data of

power transformers. It has been demonstrated that the SOM is capable of unearthing the inherent characteristics of the DGA data and presenting these revealed features in a discernible and comprehensible format. In addition, it has been validated that these revealed features can be associated with several conditions of power transformers. More importantly, through a comparison with conventional DGA interpretation schemes, it has been demonstrated that the interpretation as provided by the proposed approach is more accurate and convincing, thereby giving confidence in the technique developed. Furthermore, the proposed approach has an added advantage of allowing the "visualisation" of the evolution of health and condition of a power transformer, based on the movement of its DGA trajectory on the u-matrix illustration where various fault regions have been previously identified. In essence, the study conducted on the DGA data of power transformers forms the largest proportion and importance in this research project; various approaches proposed in this part of the study have also been applied for the analysis and interpretation of the DGA data of transformer bushings as described below.

10.2.5 Data Mining on the DGA data of Transformer Bushings

The proposed approach from the foregoing study on the DGA data of power transformers has also been applied for the DGA data of their bushings. Here again, the SOM has been found to be capable of unearthing the inherent characteristics of the DGA data and presenting these revealed features in a discernible and comprehensible format. In addition, it has been verified that these revealed features can be associated with several conditions of transformer bushings. More importantly, the proposed approach has been demonstrated to be capable of providing an easy means of fault diagnosis and fault detection in bushings of power transformers before any suspected faulty bushing is replaced.

10.2.6 Data Mining on the Multiple Sensor Data of Power Transformer

The proposed approach has also been applied for the analysis on a set of on-line monitoring data comprising of hourly measurements of various sensor variables

collected from a unit of power transformer. In this respect, it has been demonstrated that the SOM is able to unearth the inherent characteristics of the multiple sensor data and present these revealed features in a discernible and comprehensible format. More importantly, it has been illustrated that these revealed features can be applied for hourly on-line monitoring of the operating condition of the power transformer.

10.3 Future Work

As mentioned before, the proposed approach is both generic and robust and can thus be applied for the analysis and interpretation of various condition assessment data. Some of the possible future work of this research project is listed below:

- Application of the proposed approach on the DGA data of other types of power transformers such as those in the field of generation and distribution.
- Application of the proposed approach on the DGA data of other oil-filled electrical devices such as current transformers, voltage transformers, underground cables, selectors, switch-gears and circuit-breakers.
- Application of the proposed approach on the multiple sensor data of a power transformer which has been monitored on-line for a considerably longer period of time.
- Development of the proposed approach into a generic software package which can provide automatic training and selection of the optimum SOM based on a set of data.
- Assess the possibility of using correlation, regression and other statistical methods to extract useful features from the data in the pre-processing stages, so as to compress and reduce the size of the training data for SOM.
- Incorporation of AI techniques such as expert system in the post-processing stages for automatic reasoning of the extracted features or patterns from the data.

BIBLIOGRAPHY

[1] SPARLING, B.: "Managing the Life of Transformers", *RCM for Substation, Transmission and Distribution Conference*, Feb. 1999.

[2] CIGRÉ WORKGROUP 12-05: "An International Survey on Failures in Large Power Transformers", *Électra*, 1983, No. 88, pp. 21-48.

[3] WILLIAM, M. F.: "Handbook of Transformer Applications" (McGraw-Hill, 1986).

[4] ASCHWANDEN, TH., HÄSSIG, M., FUHR, J., LORIN, P., DER HOUHANESSIAN, V., ZAENGL, W., SCHENK, A., ZWEIACKER, P., PIRAS, A.: "Development and Application of New Condition Assessment Methods for Power Transformers", *International Conference on Large High Voltage Electric Systems*, 1998, Vol. 2, pp. 12.207.1-12.207.10.

[5] LEIBFRIED, T., KNORR, W., VIERECK, K., DOHNAL, D., KOSMATA, A., SUNDERMANN, U., BREITENBAUCH, B.: "On-Line Monitoring of Power Transformers – Trends, New Developments and First Experiences", *International Conference on Large High Voltage Electric Systems*, 1998, Vol. 2, pp. 12.211.1-12.211.9.

[6] KANG, P., BIRTWHISTLE, D.: "Condition Monitoring of Power Transformer On-Load Tap-Changers. I. Automatic Condition Diagnostics", *IEE Proceedings – Generation, Transmission and Distribution*, Jul. 2001, Vol. 148, pp. 301-306.

[7] KANG, P., BIRTWHISTLE, D.: "Condition Monitoring of Power Transformer On-Load Tap-Changers. II. Detection of Ageing from Vibration Signatures", *IEE Proceedings – Generation, Transmission and Distribution*, Jul. 2001, Vol. 148, pp. 307-311.

[8] DE, A., CHATTERJEE, N.: "Impulse Fault Diagnosis in Power Transformers Using Self-Organising Map and Learning Vector Quantisation", *IEE Proceedings – Generation, Transmission and Distribution*, Sept. 2001, Vol. 148, pp. 397-405.

[9] DE, A., CHATTERJEE, N.: "Recognition of Impulse Fault Patterns in Transformers Using Kohonen's Self-Organizing Feature Map", *IEEE Transactions on Power Delivery*, Apr. 2002, Vol. 17, pp. 489-494.

[10] HONG, DU, INUI, M., OHKITA, M., FUJIMURA, K., TOKUTAKA, H.: "Short-Term Prediction of Oil Temperature Change of an Indoor Transformer by Self-Organising Map (SOM)", *Proceedings of the 2002 IEEE Power Engineering Society Winter Meeting*, 2002, Vol. 2, pp. 1366-1371.

[11] WERLE, P., AKBARI, A., BORSI, H., GOCKENBACH, E.: "Partial Discharge Localisation on Power Transformers Using Neural Networks Combined with Sectional Winding Transfer Functions as Knowledge Base", *Proceedings of 2001 International Symposium on Electrical Insulating Materials*, 2001, pp. 579-582.

[12] ROIZMAN, O., DAVYDOV, V.: "Neuro-Fuzzy Computing for Large Power Transformers Monitoring and Diagnostics", *Proceedings of 18th International Conference of the North American Fuzzy Information Processing Society*, 1999, pp. 248-252.

[13] BOOTH, C., MCDONALD, J. R.: "The Use of Artificial Neural Networks for Condition Monitoring of Electrical Power Transformers", *Neurocomputing*, Dec. 1998, Vol. 23, No. 1-3, pp. 97-109.

[14] WANG, Z., LIY, Y., GRIFFIN, P. J.: "Neural Net and Expert System Diagnose Transformer Faults", *IEEE Computer Applications in Power*, Jan. 2000, Vol. 13, No. 1, pp. 50-55.

[15] DÖRNENBURG, E., STRITTMATTER, W.: "Monitoring Oil Cooled Transformers by Gas Analysis", *Brown Boveri Review*, May 1974, Vol. 61, No. 5, pp. 238-247.

[16] ROGERS, R. R.: "IEEE and IEC Codes to Interpret Incipient Faults in Transformers, Using Gas in Oil Analysis", *IEEE Transactions on Electrical Insulation*, Oct. 1978, Vol. EI-13, No. 5, pp. 349-354.

[17] DUVAL, M.: "Dissolved Gas Analysis: It Can Save Your Transformer", *IEEE Electrical Insulation Magazine*, Nov.-Dec. 1989, Vol. 5, No. 6, pp.22-27.

[18] DUVAL, M., *CIGRÉ Symposium*, 1993, pp. 110-114.

[19] IEC STANDARD – PUBLICATION 599: "Interpretation of the Analysis of Gases in Transformers and Other Oil-Filled Electrical Equipment in Service" (International Electrotechnical Commission, 1978).

[20] IEC STANDARD 60599: "Mineral Oil-Impregnated Electrical Equipment in Service – Guide to the Interpretation of Dissolved and Free Gases Analysis" (International Electrotechnical Commission, 1999).

[21] MÖLLMANN, A., PAHLAVANPOUR, B.: "New Guidelines for Interpretation of Dissolved Gas Analysis in Oil-filled Transformers", *Électra*, Oct. 1999, No. 186, pp. 31-51.

[22] CHECKSFIELD, M., WESTLAKE, A.: "Experiences with Operating and Monitoring Generator Transformers", *Proceedings of the IEE Colloquium on Transformer Life Management*, 1998, pp. 4/1-4/7.

[23] LAPWORTH, J., MCGRAIL, T.: "Transformer Failure Modes and Planned Replacement", *Proceedings of the IEE Colloquium on Transformer Life Management*, 1998, pp. 9/1-9/7.

[24] IEC STANDARD – PUBLICATION 567: "Guide for the Sampling of Gases and of Oil-filled Electrical Equipment and for the Analysis of Free and Dissolved Gases" (International Electrotechnical Commission, 1992).

[25] PAKONEN, P., LEHTIO, A., KANNUS, K., LAKERVI, E.: "RFI Measurements as a Diagnostic Technique for Medium Voltage Transformers", *International Conference on Partial Discharge*, 1993, pp. 115-116.

[26] URBANI, G. M., BROOKS, R. S.: "Using the Recovery Voltage Method to Evaluate Aging in Oil-Paper Insulation", *Proceedings of the 1998 IEEE 6th International Conference on Conduction and Breakdown in Solid Dielectrics*, 1998, pp. 93-97.

[27] STIGANT, S. A., FRANKLIN, A. C.: "The J & P Transformer Book" (John Wiley & Sons, 1973), pp. 269-270.

[28] BENGTSSON, C.: "Status and Trends in Transformer Monitoring", *IEEE Transactions on Power Delivery*, Jul. 1996, Vol. 11, No. 3, pp. 1379-1384.

[29] REASON, J.: "On-line transformer monitoring", *Electrical World*, Oct. 1995, Vol. 209, No. 10, pp. 19-26.

[30] ABB BROCHURE: "A New Monitoring Concept for On-Line Transformer Diagnostics: ABB T-Monitor".

[31] HALSTEAD, W. D.: "A Thermodynamic Assessment of the Formation of Gaseous Hydrocarbons in Faulty Transformers", *J. Inst. Petroleum*, Sept. 1973, pp. 239-241.

[32] BHATTACHARYYA, S. K., SMITH, R. E., HASKEW, T. A.: "A Neural Network Based Approach to Transformer Fault Diagnosis Using Dissolved Gas Analysis Data", *Proceedings of the 1993 NAPS*, Oct. 1993, pp. 125-129.

[33] ZHANG, Y., DING, X., LIU, Y., GRIFFIN, P. J.: "An Artificial Neural Network Approach to Transformer Fault Diagnosis", *IEEE Transactions on Power Delivery*, Oct. 1996, Vol. 11, No. 4, pp. 1836-1841.

[34] WANG, Z., ZHANG, Y., LI, C., LIU, Y.: "ANN based Transformer Fault Diagnosis", *Proceedings of the American Power Conference*, 1997, Vol. 59-1, pp. 428-432.

[35] TU, Y. M., HUANG, J. M., GAO, N., ZHU, Z. S., YAN, Z.: "Transformer Insulation Diagnosis based on Improved ANN Analysis", *Proceedings of the IEEE International Conference on Properties and Applications of Dielectric Materials*, May 1997, Vol. 1, pp. 263-266.

[36] VANEGAS, O., MIZUNO, Y., NAITO, K., KAMIYA, T.: "Diagnosis of Oil-Insulated Power Apparatus by using Neural Network Simulation", *IEEE Transactions on Dielectrics and Electrical Insulation*, Jun. 1997, Vol. 4, No. 3, pp. 290-299.

[37] GAO, W., QIAN, Z., YAN, Z.: "A Multi-Resolution System Approach to Power Transformer Insulation Diagnosis", *Proceedings of the International Symposium on Electrical Insulating Materials*, Sept. 1998, pp. 685-688.

[38] ESP, D. G., CARRILLO, M., MCGRAIL, A. J.: "Data Mining Applied to Transformer Oil Analysis Data", *Conference Record of IEEE International Symposium on Electrical Insulation*, Jun. 1998, Vol. 1, pp. 12-15.

[39] ESP, D. G., CARRILLO, M., MCGRAIL, A. J.: "Exploratory Mining of Transformer Gas-In-Oil Data", *Proceedings of 33rd Universities Power Engineering Conference*, Sept. 1998, Vol. 1, pp. 278-281.

[40] ESP, D. G., MCGRAIL, A. J.: "Data Mining Applied to Transformer Oil Analysis Data", *IEE Colloquium Insulating Liquids*, May 1999, pp. 9/1-9/7.

[41] ESP, D. G., MCGRAIL, A. J.: "Advances in Data Mining for Dissolved Gas Analysis", *Conference Records of the 2000 IEEE International Symposium on Electrical Insulation*, Apr. 2000, pp. 456-459.

[42] YANG, H. T., HUANG, Y. C.: "Intelligent Decision Support for Diagnosis of Incipient Transformer Faults using Self-Organising Polynomial Networks", *IEEE Transactions on Power System*, Aug. 1998, Vol. 13, No. 3, pp. 946-952.

[43] LIN, C. E., LING, J. M., HUANG, C. L.: "An Expert System for Transformer Fault Diagnosis Using Dissolved Gas Analysis", *IEEE Transactions on Power Delivery*, Jan. 1993, Vol. 8, No. 1, pp. 231-238.

[44] TOMSOVIC, K., TAPPER, M., INGVARSSON, T.: "Fuzzy Information Approach to Integrating Different Transformer Diagnostic Methods", *IEEE Transactions on Power Delivery*, Jul. 1993, Vol. 8, No. 3, pp. 1638-1645.

[45] HUANG, Y. C., YANG, H. T., HUANG C. L.: "Developing a New Transformer Fault Diagnosis System through Evolutionary Fuzzy Logic", *IEEE Transactions on Power Delivery*, Apr. 1997, Vol. 12, No. 2, pp. 761-767.

[46] TOMSOVIC, K., AMAR,A.: "On Refining Equipment Condition Monitoring Using Fuzzy Sets and Artificial Neural Nets", *International Journal of Engineering Intelligent Systems for Electrical Engineering and Communications*, Mar. 1997, Vol. 5, No. 1, pp. 43-50.

[47] WANG, Z., LIU, Y., GRIFFIN, P. J.: "A Combined ANN and Expert System Tool for Transformer Fault Diagnosis", *IEEE Transactions on Power Delivery*, Oct. 1998, Vol. 13, No. 4, pp. 1224-1229.

[48] IEEE STANDARD C57.104: "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers" (Institute of Electrical and Electronic Engineers, 1991).

[49] SAMMON, J. W.: "A Non-Linear Mapping for Data Structure Analysis", *IEEE Transactions on Computers*, May 1969, Vol. C-18, No. 5, pp. 401-409.

[50] KOHONEN, T.: "Self-Organizing Maps (2^{nd} Edition)" (Springer-Verlag, 1997).

[51] FAYYAD, U. M.: "Data Mining and Knowledge Discovery: Making Sense Out of Data", *IEEE Intelligent Systems & Their Applications*, Oct. 1996, Vol. 11, No. 5, pp. 20-25.

[52] FAYYAD, U. M., PIATETSKY-SHAPIRO, G., SMYTH, P.: "From Data Mining to Knowledge Discovery: An Overview" (MIT Press, 1996).

[53] CLEMENTINE (http://spss.com/spssbi/clementine/).

[54] DATAENGINE (http://www.dataengine.de/english/sp/index.htm).

[55] DATASCOPE (http://www.cygron.com/DataMining.html).

[56] IRIS EXPLORER (http://www.nag.co.uk/Welcome_IEC.html).

[57] Documentation for Data Mining Software: (http://www.andypryke.com/university/software.html).

[58] ULTSCH, A.: "Knowledge Acquisition with Self-Organizing Neural Networks", *Proceedings of the 1992 International Conference*, 1992, Vol. 1, pp. 735-739.

[59] ULTSCH, A.: "Knowledge Extraction from Self-Organizing Neural Networks", *Information and Classification*, 1993, pp. 301-306.

[60] ULTSCH, A.: "Self-Organizing Neural Networks for Visualisation and Classification", *Information and Classification*, 1993, pp. 307-313.

[61] SIMULA, O., ALHONIEMI, E., HOLLMEN, J., VESANTO, J.: "Analysis of Complex Systems using the Self-Organizing Map", *Proceedings of the 1997 International Conference on Neural Information Processing and Intelligent Information Systems*, 1998, Vol. 2, pp.1313-1317.

[62] VESANTO, J., VASARA, P., HELMINEN, R. R., SIMULA, O.: "Integrating Environmental, Technological and Financial Data in Forest Industry Analysis", *Neural Networks: Best Practice in Europe*, 1997, pp. 153-156.

[63] VESANTO, J., AHOLA, J.: "Hunting for Correlations in Data Using the Self-Organizing Map Units", *Proceedings of the International ICSC Congress on Computational Intelligence Methods and Applications*, 1999, pp. 7-13.

[64] VESANTO, J.: "Data Mining Techniques Based on the Self-Organizing Map", *Thesis for the degree of Master of Science in Engineering*, Department of Engineering Physics and Mathematics, Helsinki University of Technology, 26th May 1997.

[65] Neural Network Research Centre, Helsinki University of Technology: "On-line Documentation for SOM Toolbox V2.0 Beta for Matlab 5.0", (http://www.cis.hut.fi/projects/somtoolbox/package/docs2/somtoolbox.html).

[66] SOM_PAK (http://www.cis.hut.fi/research/som_lvq_pak.shtml).

[67] SOM_TOOLBOX (http://www.cis.hut.fi/projects/somtoolbox/).

APPENDIX

SCALING METHODS FOR TRAINING DATA

Three scaling methods were considered for the pre-processing of input data before submitted for the training of self-organising map (SOM), i.e. the "range", "variance" and "logarithmic" methods. These scaling methods are briefly explained in the following sections.

1. The "Range" Scaling Method

The "range" method transforms the minimum and maximum values of each component of the input data into "0" and "1" respectively; all other values in each component of the input data are scaled according to Equation (1).

$$x_{new} = \frac{x_{old} - min}{max - min} \quad \text{.....(1)}$$

Where, x_{new} and x_{old} are the scaled and original values respectively. And, *min* and *max* are the minimum and maximum values of the un-scaled data.

2. The "Variance" Scaling Method

The "variance" method transforms the values in each component of the input data such that the mean and variance of each component of the scaled input data are "0" and "1" respectively, as shown in Equation (2).

$$x_{new} = \frac{x_{old} - mean}{stdev} \quad \text{.....(2)}$$

Where, x_{new} and x_{old} are the scaled and original values respectively. And, *mean* and *stdev* are the mean and standard deviation of the un-scaled data.

3. The "Logarithm" Scaling Method

The "logarithm" method transforms the values in each component of the input data according to Equation (3). It shall be noted that natural logarithm has been used in this case and the offset of "1" has been added to the values in each component of the input data in order to prevent the occurrence of errors due to the presence of zeros in the data.

$$x_{new} = ln(1 + x_{old}) \quad \text{.....(3)}$$

Where, x_{new} and x_{old} are the scaled and original values respectively.

RELATED PUBLICATIONS

[1] THANG, K. F., AGGARWAL, R. K., ESP, D. G., MCGRAIL, A. J.: "Statistical and Neural Network Analysis of Dissolved Gases in Power Transformers", *Proceedings of Eight International Conference on Dielectric Materials, Measurements and Applications*, Sept. 2000, IEE Conference Publication No. 473, pp. 324-329.

[2] THANG, K. F., AGGARWAL, R. K., MCGRAIL, A. J., ESP, D. G.: "Application of Self-Organising Map Algorithm for Analysis and Interpretation of Dissolved Gases in Power Transformers", *Proceedings of 2001 IEEE Power Engineering Society Summer Meeting*, July 2001, Paper No. 0-7803-7031-7/01.

[3] THANG, K. F., AGGARWAL, R. K., MCGRAIL, A. J., ESP, D. G.: "Data Mining Approach for Analysis of Power Transformer Dissolved Gas Records Using the Self-Organising Map", Submitted to IEEE Transactions on Power Delivery.

STATISTICAL AND NEURAL NETWORK ANALYSIS OF DISSOLVED GASES IN POWER TRANSFORMERS

K F Thang*, R K Aggarwal*, D G Esp+, A J McGrail+

University of Bath*, UK, The National Grid Company plc+, UK

ABSTRACT

The onset of electrical discharges or thermal stresses in mineral oil or cellulose insulation of a power transformer can cause the degradation of these materials with the formation of various dissolved gases. These dissolved gases can be extracted and identified with the application of gas chromatography. The overall process, from oil sampling to gas identification, is known as dissolved gas analysis (DGA). In this paper, a comparison of conventional DGA interpretation schemes is briefly presented. Moreover, some new artificial intelligence (AI) techniques for transformer incipient fault diagnosis based on DGA data, are also discussed. The second part of the paper reports on the initial work performed for the proposed new approach. This includes simple statistical analysis on DGA records and is followed by high-level data-mining (DM) using self-organising map (SOM) algorithm. The inherent data 'structure' revealed from the latter part of the analysis could hypothetically be associated with certain transformer faults, either electrical, thermal or cellulose decomposition. The proposed approach could provide a viable alternative for transformer incipient fault diagnosis and condition-monitoring applications.

INTRODUCTION

The use of DGA for transformer incipient fault diagnosis is verified by the fact that insulation materials within the transformer tank, such as mineral oil and paper pressboard etc, will disintegrate when subjected to a high degree of temperature as result of faults. The decomposed products are mainly in the form of gases, which will eventually dissolve in the tank oil or evolve to the gas-collecting relay. The types and quantity of dissolved gases depend fundamentally on the degree of decomposition temperature, or more particularly on the amount of energy available to decompose the insulation materials.

As discussed in IEC 60599, partial discharge occurs in the case of low-level energy, such as breakdown in gas-filled cavities resulting from incomplete oil-impregnation. In this case, the major gas produced is hydrogen (H_2). In other types of faults, the decomposition of oil is mainly caused by heat. Decomposition occurs at normal operating temperature, producing mainly hydrogen (H_2) and methane (CH_4). Higher decomposition temperature, resulting from hot spots or conductor overheating, produces mainly methane (CH_4). With further increases in decomposition temperature, an increasing amount of ethane (C_2H_6) and ethylene (C_2H_4) will be liberated. In the case of a much higher temperature resulting from disruptive electrical faults, such as sparking or flashover, the production of acetylene (C_2H_2) becomes significant. On the other hand, carbon dioxide (CO_2) and carbon monoxide (CO) are produced as a result of decomposition of cellulose insulation, which may be initiated by either electrical or thermal faults.

Owing to the credible association of dissolved gases with types of faults occurring within a power transformer, various DGA interpretation schemes have been established, such as Dörnenburg's ratios, Rogers' ratios, IEC's ratios, CIGRÉ's method and Duval's Triangle. The application of these conventional schemes for fault diagnosis is less than perfect since quite often an interpretation is not possible and the use of different scheme will result in different fault diagnoses. Consequently, attempts have been made to employ artificial intelligence (AI) techniques to tackle these drawbacks. Although these new approaches are still dependent on conventional interpretation schemes for their development, significant improvement over conventional schemes have been reported.

A new approach for transformer fault diagnosis and condition monitoring is proposed in this paper. It is based on a type of unsupervised neural-network known as self-organising map (SOM) algorithm, as will be discussed in the following sections.

COMPARISON ON CONVENTIONAL DGA INTERPRETATION SCHEMES

Some commonly applied DGA interpretation schemes are compared as follows:

Dörnenburg ratios. Dörnenburg's method is based on ratios CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/C_2H_2 and C_2H_2/CH_4. Three types of faults are detectable, i.e. local overheating (thermal fault), partial discharge and other types of discharges (electrical fault). Dörnenburg recommendation on the application of this method

should commence with determining whether dissolved gases (including CO and CO_2) are below the quoted concentration limits; faults are suspected if one or more gases exceed the limits. Rules were also outlined to determine the applicability of these ratios. Implementation of Dörnenburg's method may result in a significant number of 'no-interpretation' cases arising from incompleteness in the ratio-ranges and non-applicability of the method.

Rogers' ratios. The Rogers' method utilises four ratios, viz. CH_4/H_2, C_2H_6/CH_4, C_2H_4/C_2H_6 and C_2H_2/C_2H_4. Diagnosis of faults is accomplished via a simple coding scheme based on ranges of ratios. Four conditions are detectable, i.e. normal ageing, partial discharge with or without tracking, thermal fault and electrical fault of various degrees of severity. It is better than Dörnenburg's method since a broader range of ratio-combinations is included and hence leads to a significant reduction of 'no-interpretation' cases. Nevertheless, no consideration is given for dissolved gases below 'normal' concentration values. Therefore, exact implementation of Rogers' method may lead to many mis-interpreted cases.

Duval's triangle. Duval's method is special since fault diagnosis is performed based on visualisation of the location of dissolved gases in the triangular map. Only three dissolved gases are needed, viz. CH_4, C_2H_4 and C_2H_2. These gases must be transformed into triangular co-ordinates before being plotted onto the triangle. Generally, three types of faults are detectable, i.e. partial discharge, high and low energy arcing (electrical fault) and hot spots of various temperature ranges (thermal fault). Since no region is designated for normal ageing condition, careless implementation of Duval's triangle will result in the diagnosis of either one of the three faults. In view of this problem, dissolved gases should be assessed for their 'normality' before being interpreted using Duval's triangle.

IEC's Ratios (IEC 60599). Fault diagnosis scheme recommended by IEC originated from Rogers' method, except that the ratio C_2H_6/CH_4 was dropped since it only indicated a limited temperature range of decomposition. Four conditions are detectable, i.e. normal ageing, partial discharge of low and high energy density, thermal faults and electrical faults of various degrees of severity. However, no attempt is made to identify both thermal and electrical faults into more precise subtypes. The first edition of IEC's method (IEC 60599-1978) is based on simple coding scheme while the second edition (IEC 60599 –1999) utilises the revised ratio ranges directly. Assessment of dissolved gases for 'normality' limits is required before being interpreted using ratios. Other improvement in the second edition of IEC method is the use of 3D graphical representation for the ratio ranges. Those cases where diagnosis of fault is not possible can be plotted onto the graph and its nearest distance to a certain fault region can then be observed.

CIGRÉ's method. CIGRÉ method utilises both key gas ratios and key gas concentrations for fault diagnosis. The key gas ratios are C_2H_2/C_2H_6, H_2/CH_4, C_2H_4/C_2H_6, C_2H_2/H_2 and CO_2/CO and the key gas concentrations are C_2H_2, H_2, sum of carbon hydrides, CO and CO_2. Ratio ranges and concentration limits are suggested for both key gas ratios and key gas concentrations; a fault is suspected if any one of the key gas ratio or key gas concentration exceeds the limit. Clear guidelines are not given concerning the combined application of both methods. Moreover, the suggested limit of 10000 for CO is less than convincing. However, a transformer is considered healthy if all ratios and concentrations are below the limits. Nevertheless, the use of CIGRÉ method is beneficial since it allows simultaneous detection of two or more faults.

Correct and 'practical' implementations of the above schemes require the availability of normal concentration values for dissolved gases. If a large DGA database is available, 90%-typical concentration values are suggested by IEC as the normality values. Other transformer information such as transformer type, age and operating condition must also be taken into consideration when making repair decision based on resultant fault diagnosis. In conclusion, currently available DGA interpretation schemes are far from satisfactory and comprehensive. Consequently, fault diagnosis based on DGA is still heavily dependent on experiences of human experts.

REVIEW OF NEW AI APPROACHES FOR TRANSFORMER FAULT DIAGNOSIS

The conventional DGA interpretation schemes have been known to suffer from the following drawbacks:

(1) Gas ratios or methods defined by these schemes are mainly developed based on human judgement. No systematic attempt has been make to actually 'learn' from the measured DGA data.
(2) There is still a high degree of inconsistency and ambiguity when applying these schemes, owing to incompleteness of the possible ratio-combinations and doubts on the validity of the defined ratio ranges.
(3) These schemes are still unable to detect with high confidence multiple faults that occur concurrently within the transformer.
(4) These schemes are unable to detect new or unknown faults owing to lack of expert knowledge in them.

Consequently, attempts have been made to utilise artificial intelligence (AI) techniques to tackle these drawbacks. There are two major categories of AI

approaches. The first category attempts to learn the inherent correlation between actual transformer faults and its corresponding DGA records. An Example of this category is the artificial neural network (ANN). The second category of AI approaches is dependent on conventional schemes for its development, but the incompleteness of the conventional schemes is addressed by incorporating a fuzzy logic (FL) algorithm. An example of this category is fuzzy expert system (FES). Variants of these two categories exist as well.

Interest in applying ANN for transformer fault diagnosis based on DGA started with Bhattacharyya et al (1) in 1993. A total of eleven gas inputs were applied to a supervised ANN, with the training data consisting of DGA records with actual observed faults. It has been reported that the use of this approach has resulted in higher diagnosis accuracy when compared with Rogers' and Dörnenburg's methods. However, only electrical and thermal faults are detectable. The application of ANN has been advanced further by Zhang et al (2), with the development of two supervised ANNs were developed. The occurrences of normal condition, overheating, corona and arcing are detectable using one ANN and the involvement of cellulose decomposition in fault is detectable using the other ANN.

By comparison, the use of FES is more acceptable by DGA experts, by virtue of the fact that this approach allows their experiences and knowledge to be incorporated. Although its development is still dependent on conventional DGA interpretation schemes, the uncertainties of normality thresholds and ratio ranges have been tackled with the use of fuzzy logic. An example of such approach was reported by Lin et al (3) and. Latter in 1998, and an evolutionary-programming approach was introduced by Huang et al (4) to further adjust the rules and membership functions in the FES. The most promising AI approach to date is the combined ANN and expert system (ES) method developed by Wang et al (5). It has the advantage of utilising ANN to learn from actual DGA records. Furthermore, it is also able to incorporate expert knowledge and conventional DGA interpretation schemes. The final diagnosis is produced based on the combination of both ANN and ES methods.

Notice that all currently available AI approaches require either DGA records with actual observed faults or conventional interpretation schemes for their developments. Acquirement of DGA records with actual faults observed in transformers has been known to be difficult, since it is costly to disconnect a particular transformer just for the purposes of confirming a suspected fault. Therefore, the supervised ANN approach will not be able to detect 'all' possible faults due to the lack of 'good' fault samples for its training. On the other hand, the dependency of FES on conventional DGA interpretation schemes has also lead to its inheritance of the inherent drawbacks of conventional schemes, as discussed previously.

STATISTICAL ANALYSIS ON DGA DATA

Large DGA database is available for analysis, courtesy of the National Grid Company plc, UK. Association has been made between the amount of DGA data and the number of transformers and transformer-age for different voltage ratios and power ratings, as illustrated in Figures 1, 2, 3 and 4, respectively

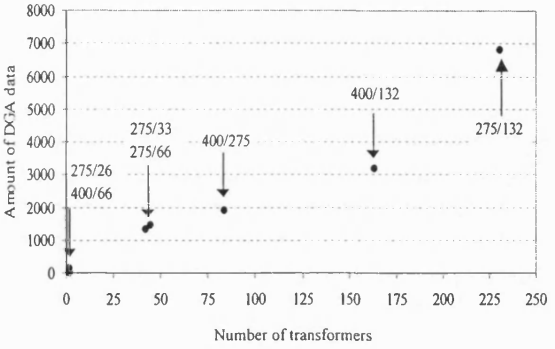


FIGURE 1: Amount of DGA data versus number of transformers for different voltage ratios.

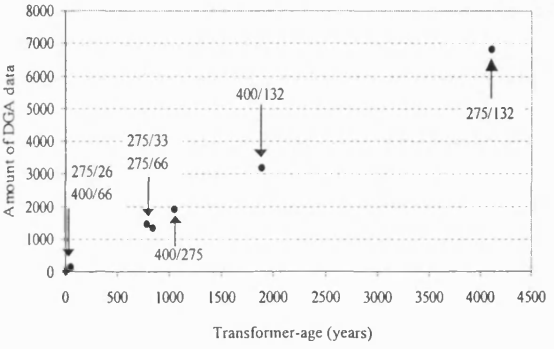


FIGURE 2: Amount of DGA data versus transformer-age for different voltage ratios.

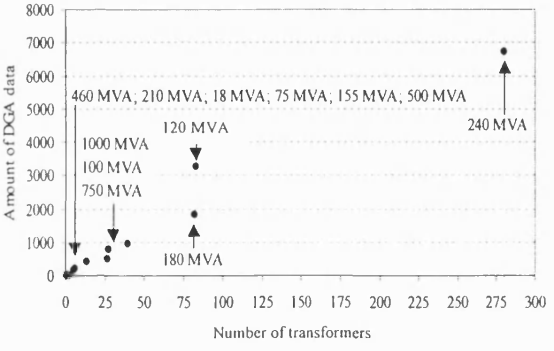


FIGURE 3: Amount of DGA data versus number of transformers for different power ratings.

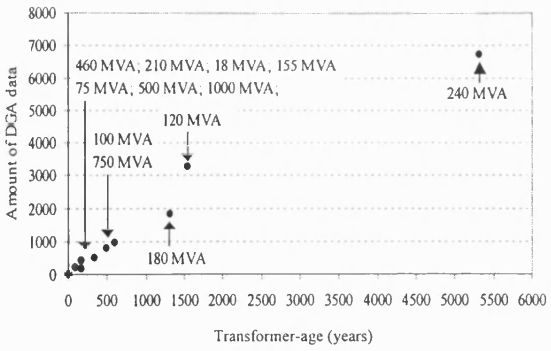


FIGURE 4: Amount of DGA data versus transformer-age for different power ratings.

As observed from above figures, there is close agreement between the total number of transformers and transformer-ages for different voltage ratios and power ratings. It was also discovered that the largest number of transformers is associated with voltage ratio of 275/132 and power rating of 240 MVA.

The 90% typical concentration values for dissolved gases were also computed from the database, as shown in Table 1:

Dissolved gases	90% typical values (µl/l)
Moisture (H_2O)	19
Nitrogen (N_2)	70000
Oxygen (O_2)	25200
Carbon dioxide (CO_2)	4260
Carbon monoxide (CO)	520
Hydrogen (H_2)	85
Methane (CH_4)	47
Ethane (C_2H_6)	31
Ethylene (C_2H_4)	53
Acetylene (C_2H_2)	13

TABLE 1 - 90% typical concentration values for dissolved gases

In addition, other statistical analyses such as frequency distributions and scatter plots were also examined. Interesting features were obtained from the foregoing analysis.

NEURAL NETWORK ANALYSIS ON DGA DATA

A new approach for transformer incipient fault diagnosis and condition monitoring has been proposed herein. It is essentially based on a type of unsupervised neural network known as self-organising map (SOM) algorithm. The SOM is capable of learning from data and displaying the data 'structure' in a visually discernible format. The data 'structure' refers to inherent organisation of data that can only be displayed effectively through a high-level data-analysis tool. The SOM is comprised of a lattice of neurons arranged in hexagonal configuration. Desired outputs are not needed since weight vectors of neurons are updated through an iterative process of associating each training data to each neuron. Hence, only 'raw' DGA records are needed for the training of SOM.

A DGA dataset was compiled for the initial feasibility study of the proposed approach, which consists of 755 DGA records from transformers with power rating of 240 MVA and voltage ratio of 275/132 and from three different manufacturers. A 20 units by 30 units SOM was used for the training. Training inputs were formed by 755 DGA records with each of them consisting of 7 dissolved gases, viz. CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2. Each dissolved gas was normalised to the range of 0 to 1 so as to prevent the masking of larger-valued gases such as CO_2 and CO on smaller-valued gases such as C_2H_2. Training was performed in two phases: rough-ordering and fine-tuning.

The trained SOM was visualised using U-matrix representation, as shown in the first map of Figure 5. The organisation of each dissolved gas can be displayed in the form of component plane, also illustrated in Figure 5. Essentially, the U-matrix representation is a summarised illustration of all the component planes.

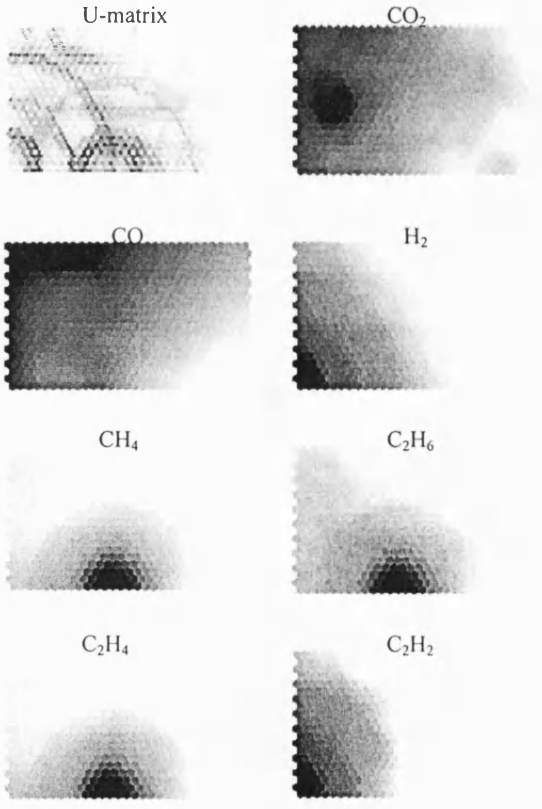


FIGURE 5 – Visualisation of SOM.

The U-matrix illustration essentially depicts the inherent 'structure' of the DGA data. The colour-depth of a 64-level grey scale as displayed in the U-matrix map is used to represent the 'euclidean distance' between one neuron to the other. Therefore, the revealed data 'structure' can be easily identified by means of the black-border around a lighter-coloured region. All corresponding data vectors located within a bordered region are said to be similar to one another in the vector space. In contrast, the colour-depth depicted in component plane of each dissolved gas represents the magnitude. The deeper the colour, the larger is the magnitude of each dissolved gas.

As observed from the U-matrix illustration in Figure 5, there are three part-circular clusters in the U-matrix map. The bottom-left cluster, which comprises of three co-centre quarter-circulars, corresponds to regions where H_2 and C_2H_2 have a significant presence. On the other hand, the bottom-middle cluster, which comprises of two co-centred half-circulars, coincides with regions where CH_4, C_2H_6 and C_2H_4 have significant magnitudes. These two major clusters also intercept with each other, forming yet another two sub-clusters. In addition, a half-circular cluster is also observed in the upper-middle region and it corresponds to a region where CO has the largest magnitude.

Thus, the first-step in interpreting the map was to classify it into 9 regions, as illustrated in Figure 6. DGA training data corresponding to each region was then obtained and analysed.

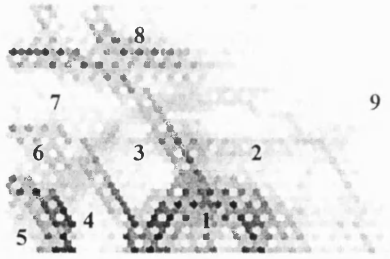


FIGURE 6 – Division of analytical regions

It was discovered that region 1 is noticeably higher only in C_2H_4, which is the key gas for indicating thermal abnormalities. Moreover, both CH_4 and C_2H_6 are the highest among all regions as well. Therefore, region 1 is hypothetically designated as thermal fault of degree of severity 2 (TF-2). Relatively high concentrations of these three gases were observed in region 2 as well, but still lower than that of region 1. Thus, region 2 is hypothetically designated as thermal fault of degree of severity 1 (TF-1). Significant amount of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 were also found for both regions 3 and 4. However, concentrations of H_2 and C_2H_2 are noticeably higher than that of region 1 and 2. Hence, regions 3 and 4 were hypothetically labelled as electrical fault and thermal fault of degree of severity 1 and 2 (EF_TF-1 and EF_TF-2) respectively. On the other hand, region 5 was typified by the largest concentrations of C_2H_2 and H_2; it is therefore hypothetically designated as an electrical fault of degree of severity 3 (EF-3). Both regions 6 and 7 are the outer clusters of region 5, and have high quantities of C_2H_2 and H_2. Since concentrations of these gases are still lower than that of region 5, region 6 and 7 can be assumed to be indicating of electrical fault of lower severity. Therefore, regions 6 and 7 are labelled as electrical fault of degree of severity 2 and 1 (EF-2 and EF-1) respectively. Region 8 was observed to be higher in CO only, and it can therefore be hypothetically designated as cellulose decomposition (CD). Finally, all dissolved gas levels in region 9 are very low; it is hence indicative of a normal ageing condition in a transformer and labelled as NORMAL hypothetically. However, no indication of partial discharge was observed in the map. The U-matrix illustration with hypothetically assigned fault regions is shown in Figure 7.

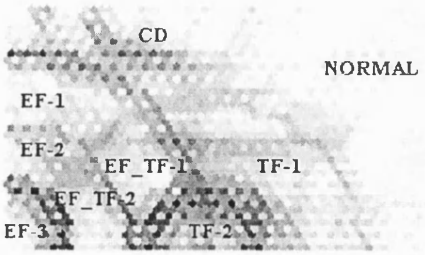


FIGURE 7 – Hypothetical identification of fault regions

In order to investigate the validity of the hypothesis in the context of conventional DGA interpretation schemes, the U-matrix illustration is overlaid with six previously described interpretation schemes. i.e. Dörnenburg's ratios, Rogers' ratios, IEC's ratios (IEC 60599: 1978 and 1999), Duval's triangle and CIGRE's method. No attempts have been made to divide the partial discharge (PD), electrical fault (EF), thermal fault (TF) and cellulose decomposition (CD) into more detail sub-types, since it is sufficient in the initial stage, for four major fault-types (EF, TF, PD and CD) to be determined. In addition, the data where conventional schemes are unable to provide interpretation were identified and labelled as 'NO INTERPRETATION'. Finally, dissolved gases were assessed for their 'normality' before being interpreted for fault. This is done as to 'mimic', as close as possible, the practical implementation of these schemes. The overlaid maps are illustrated in Figure 8.

DISCUSSIONS AND FUTURE WORK

As observed from Figure 8, fault regions as interpreted by various DGA schemes are fairly consistent with the hypothetical regions as illustrated in Figure 7. However, all but CIGRÉ's method were unable to identify simulta-

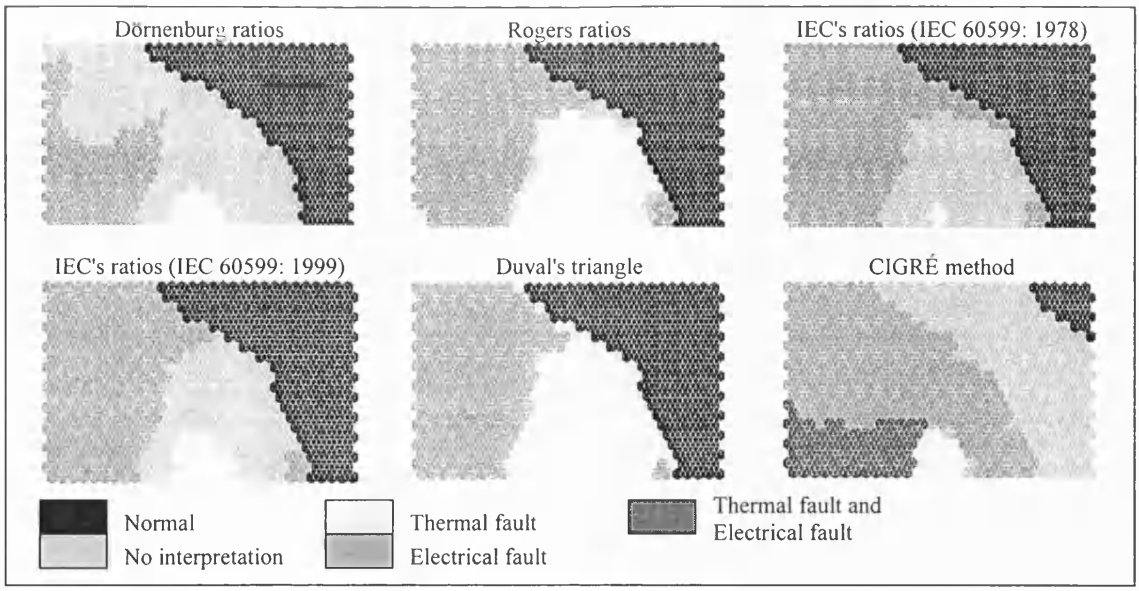


FIGURE 8: Identification of fault region based on various DGA interpretation schemes.

-neous occurrence of both electrical and thermal faults. In addition, the non-occurrence of partial discharge as concluded from previous study are also 'confirmed' by various DGA interpretation schemes. However, these schemes were unable to identify the hypothetical cellulose decomposition, therefore 'new' fault pattern has been revealed by the SOM. The 'weakness' of Dörnenburg's method is also revealed as dictacted by a large 'no-interpretation' region. Roger's method was able to detect both electrical and thermal faults very satisfactorily. Both versions of IEC' ratios were unable to detect the hypothetical thermal fault region owing to the lack of code '1 2 2' in the standard. On the other hand, Interpretation provided by Duval's method is good only if all DGA data are assessed for 'normality' before being interpretated for fault. However, diagnosis method recommended by CIGRÉ is less convincing since it labels the hypothetical less severe fault region (TF-1 in Figure 7) as electrical fault. If proven to be valid with high confidence, the hypothetical map shown in Figure 7 is more beneficial since it has been developed based on inherent 'structure' of the DGA data; no DGA data with actual faults or conventional schemes were required. Future work will include investigating DGA data of greater quantity and variety, and also proposing a credible approach to verify the hypothesis on identified fault regions.

CONCLUSIONS

Comparison on various DGA interpretation schemes have been presented. A brief review on AI approaches are also given. Finally, a new approach for transformer fault diagnosis based on DGA is proposed in the paper

ACKNOWLEDGEMENT

Assistance and contribution by the National Grid Company plc is appreciated.

REFERENCES

1. Bhattacharyya, S. K., Smith, R. E., Haskew, T. A., 1993, " A neural network based approach to transformer fault diagnosis using dissolved gas analysis", Procs. NAPS, 125-129

2. Zhang, Y., Ding, X., Liu, Y., Griffin, P. J., 1996, "An artificial neural network approach to transformer fault diagnosis", IEEE Trans. Power Delivery, 11, 1836-1841

3. Lin, C. E., Ling, J. M., Huang, C. L., 1993, "An expert system for transformer fault diagnosis using dissolved gas analysis", IEEE Trans. Power Delivery, 8, 231-238

4. Huang, Y. C., Yang, H. T., Huang C. L., 1997, "Developing a new transformer fault diagnosis system through evolutionary Fuzzy Logic", IEEE Trans. Power Delivery, 12, 761-767.

5. Wang, Z., Liu, Y., Griffin, P. J., 1998, "A combined ANN and expert system tool for transformer insulation diagnosis", IEEE Trans. Power Delivery, 13, 1224-1229

Application of Self-Organising Map Algorithm for Analysis and Interpretation of Dissolved Gases in Power Transformers

K.F. Thang
University of Bath
Bath, BA2 7AY, UK

R.K. Aggarwal
University of Bath
Bath, BA2 7AY, UK

A. J. McGrail
The National Grid Company plc
Leatherhead, KT22 7ST, UK

D.G. Esp
The National Grid Company plc
Wokingham, RG31 5BN, UK

Abstract: **Onset of incipient faults in power transformers can degrade the mineral oil and cellulose insulation, leading to the formation of dissolved gases. The process from oil sampling to quantification of gases is known as dissolved gas analysis (DGA). Despite the availability of DGA interpretation schemes and artificial intelligence (AI) methods for transformer condition monitoring (CM) based on DGA data, it is pointed out in this paper that these approaches are less than ideal and practical in implementation. In view of that, this paper illustrates a novel approach for analysis and interpretation of DGA data, which leads to a more credible CM of power transformers. The proposed approach, which is based on the self-organising map (SOM) algorithm, has been validated using real fault-cases and thereby is proven to be more reliable in portraying the current condition of power transformers.**

Keywords: **Incipient faults; Power transformer; Dissolved gas analysis; Artificial intelligence; Condition monitoring; Self-organising map.**

I. INTRODUCTION

Mineral oil in power transformers consists of many different hydrocarbon molecules. Electrical and thermal faults can break-up bonds linking these molecules, which lead to the formation of gases. For a slowly developing or incipient fault, the gases formed will dissolve in oil, with only a small proportion diffusing from the oil into any gas phase above it. This kind of fault can therefore be monitored using dissolved gas analysis (DGA).

DGA requires the sampling of seven key gases as stated in IEC 599 Standard [1], viz. carbon dioxide (CO_2), carbon monoxide (CO), hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4) and acetylene (C_2H_2). As described in IEC 599 Standard, partial discharge (PD) occurs in the case of low-level energy, such as breakdown in gas-filled cavities resulting from incomplete oil-impregnation. In this case, the major gas produced is H_2. In other types of fault, the decomposition of oil is mainly caused by heat. Decomposition occurs at normal operating temperature, producing mainly H_2 and CH_4. Higher decomposition temperature, resulting from thermal fault (TF) such as hot-spot and overheating, produces mainly CH_4. With further increases in temperature, an increasing amount of C_2H_6 and C_2H_4 will be released. In the case of a much higher temperature resulting from disruptive faults such as electrical discharges (ED), the production of C_2H_2 becomes significant. On the other hand, if cellulose materials such as paper, pressboard etc. are involved at the location of the fault, further gases, principally CO_2 and CO will also be generated.

Owing to the credible relationship between the relative composition of dissolved gases and the type and severity of faults, several DGA interpretation schemes have been established, such as the well-known Dörnenburg Ratios [2], Rogers Ratios [3], IEC Ratios [1] and Duval Triangle [4]. The implementation of these schemes, however, is less than ideal since different diagnosis of fault is often resulted. Attempts have been made to employ AI methods [5-9] for improving fault diagnosis and condition monitoring (CM) of power transformers. Although some of these approaches are still dependent on DGA schemes for their development, improvement over conventional schemes have been reported.

A novel approach for analysis and interpretation of DGA data is proposed in this paper. Unlike the above stated methods, the proposed approach does not depend on DGA interpretation schemes for its development; only measured DGA data is needed. A brief description and comparison of DGA interpretation schemes and currently available AI methods for CM of power transformers is presented in first part of the paper. The second part of the paper concentrates on introducing the self-organising map (SOM) algorithm, outlining the proposed approach, and presenting the performance and validation of the proposed approach.

II. REVIEW OF DGA SCHEMES AND AI APPROACHES FOR CM OF POWER TRANSFORMERS

A. DGA Interpretation Schemes

Several well-known DGA schemes are, for example, Dörnenburg Ratios [2], Rogers Ratios [3], IEC Ratios [1] and Duval Triangle [4]. These schemes have been implemented, either in improvised or modified format, by various power utilities throughout the world. The use of these schemes requires the computation of several key-gas ratios, as listed in Table 1. Fault diagnosis is achieved by associating these ratios with a set of pre-defined faults or conditions.

Table 1.

DGA Scheme	Key-Gas Ratios
Dörnenburg Ratios	$\frac{CH_4}{H_2}$; $\frac{C_2H_2}{C_2H_4}$; $\frac{C_2H_6}{C_2H_2}$; $\frac{C_2H_2}{CH_4}$
Rogers Ratios	$\frac{CH_4}{H_2}$; $\frac{C_2H_6}{CH_4}$; $\frac{C_2H_4}{C_2H_6}$; $\frac{C_2H_2}{C_2H_4}$
IEC Ratios	$\frac{C_2H_2}{C_2H_4}$; $\frac{CH_4}{H_2}$; $\frac{C_2H_4}{C_2H_6}$
Duval Triangle*	$\%C_2H_2$; $\%C_2H_4$; $\%CH_4$

* Note: For Duval Triangle, %[Gas] is calculated by using [Gas] as numerator and the sum of all three gases as denominator, percentage is then computed on the ratio.

Generally, three main types of faults are detectable using these schemes, viz. PD, TF and ED. Before subjecting the DGA data for interpretation, a decision has to be made on whether a fault is suspected. This is done by referring to the typical values of gas concentrations, which is supplied in Dörnenburg's method and suggested by the IEC 599 Standard. If all key-gases are below the typical values, no fault is suspected and the power transformer is regarded as operating normally. However, if any one of gases exceeds the stated typical values, a fault is suspected and hence it can be determined by calculating the key-gas ratios as shown in Table 1. In the event that a large DGA database is available, the typical values should be calculated from it based on a 90% or 95% limit [1] in order to achieve a more accurate and convincing diagnosis.

Although the aforementioned DGA interpretation schemes have been widely accepted by the power industries, they are known to have suffered from following drawbacks:

1) Gas ratios or approaches defined are mainly developed based on judgement and experience of human experts; no "systematic" attempt has been make to actually "learn" from the measured DGA records in the database.
2) There is still a high degree of inconsistency and ambiguity when applying these schemes, owing to the incompleteness of the possible ratio-combinations and uncertainty on the validity of the defined ranges of key-gas ratios.
3) The application of these schemes may sometimes result in the "no-interpretation" situations, whereby a decision on diagnosis has to be solely dependent on human expert.
4) These schemes are still unable to detect with high confidence multiple faults that occur concurrently within the power transformer.
5) These schemes are unable to detect "new" or "unknown" faults owing to a lack of expert knowledge within them.

B. AI Approaches

Attempts have been made to utilise AI techniques to address the drawbacks of DGA interpretation schemes. These AI approaches can be broadly categorised into two groups, as illustrated below.

The first category of utilises actual fault-cases in an attempt to learn the inherent relationship between the actual condition of power transformers and their corresponding DGA measurements. An example of this category is the supervised artificial neural network (ANN), as reported by Bhattacharyya *et al.* [5] and Zhang *et al.* [6]. Generally, a set of input, target-output samples is required for the training of supervised ANN. The inputs are, for example, key gases, viz. H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, and the outputs are transformer conditions, viz. PD, TF and ED, that have been identified through internal examination of faulty transformers. However, acquisition of DGA records with actual faults observed in transformers is difficult since it is costly and impractical to dismantle a transformer solely for the purposes of confirming a suspected fault. Consequently, the supervised ANN approach will not be able to detect all possible faults owing to the lack of many "genuine" fault-cases for its training.

The second category relies mainly on an expert system (ES) and fuzzy logic (FL), and the technique is commonly known as the fuzzy expert system (FES). This approach allows the incorporation of expert experience into its knowledge base, whereas its core is mainly built on several DGA interpretation schemes. The incorporation of FL is beneficial since the uncertainty of normality thresholds and ratio ranges can be effectively dealt-with. Examples of such an approach were reported by Lin *et al.* [7] and Tomsovic *et al.* [8]. However, the dependency of FES on conventional DGA schemes has led to its inheritance of some the drawbacks of these schemes, as stated in the previous section.

III. APPLICATION OF SOM FOR ANALYSIS AND INTERPRETATION OF DGA DATA

A. The SOM Algorithm

The SOM is constructed by neurons located on a regular two-dimensional grid, as illustrated in Fig. 1 [9]. Neurons are arranged in either hexagonal or rectangular configuration; the former is preferred owing to its effectiveness in visualisation. The configuration of SOM is such that it defines a mapping from the input data space $\Re^n$, where n is the dimension of input space onto a two-dimensional array of neurons. Every neurons i is associated with an n-dimensional reference vector $m_i = [\mu_{i1}, \mu_{i2}, \mu_{i3}, \ldots, \mu_{in}]^T$. An input vector $x = [\xi_1, \xi_2, \xi_3, \ldots\ldots, \xi_n]^T$ is connected to all neurons in parallel via reference values μ_{ij}, where $j = 1, \ldots, n$, which are different for every neuron.

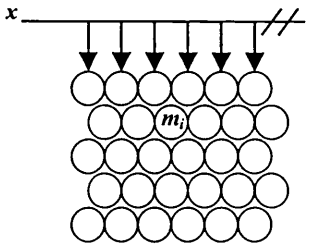


Fig 1. Neurons arranged in hexagonal lattice in a two-dimensional SOM

Once the reference vectors m_i have been initialised, the training of SOM commences by first choosing an input vector x, at time-step t, randomly from the input data set. The Euclidean distances between x and all reference vectors m_i are calculated. The reference vector that is closest to x is known as the best-matching unit (BMU), denoted as c. Equation (1) illustrates the above stated relation:

$$c = arg\,min_i\{\|x - m_i\|\} \quad (1)$$

All reference vectors are then updated, according to (2). Note that $h_{ci}(t)$ is the neighbourhood kernel around the BMU c at time-step t. It can be defined as the function of the learning rate $\alpha(t)$ and the neighbourhood function $h(\|r_c - r_i\|, t)$, as illustrated in (3). Some commonly used neighbourhood functions are, for example, the bubble function and the gaussian function [10].

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)] \quad (2)$$

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c - r_i\|, t) \quad (3)$$

If the SOM is sufficiently trained in accordance to some pre-set criterion, effective mapping from input data set to the two-dimensional reference vectors can be achieved. The inherent characteristic of data can therefore be "learned" and visualised using some data visualisation tools [10].

B. The Proposed Approach

The SOM algorithm is employed herein for the analysis and interpretation of DGA data, with the benefit of its non-dependency on actual fault-cases and DGA interpretation schemes for its modelling. In addition, the "knowledge" learned by SOM can be displayed in a comprehensible format, from which valuable information can be extracted and utilised for CM of power transformers.

A DGA data set was extracted from the full database for the feasibility study of the proposed approach. This data set, which consists of 755 measured DGA records, corresponds to power transformers constructed by three different manufacturers, and of two different voltage ratios and one power rating, as illustrated in Table 2. This data set can be regarded as a "non-ideal" representation of the entire DGA database, by virtue of the fact the diversity is reflected by DGA records of transformers from various substations, of different manufacturers, and of voltage ratios and power rating that account for over 50% of the entire transformer population.

Table 2.

Manufacturer*	Voltage Ratio	Power Rating (MVA)	Number of Samples
I	400/132	240	72
I	275/132	240	91
II	400/132	240	94
II	275/132	240	187
III	400/132	240	108
III	275/132	240	203

* Note: Roman numerics are used for the representation of manufacturers

Four input-combinations and pre-processing methods have been tested on the data set, resulting in a total of sixteen SOM configurations, as illustrated in Table 3. Note that those marked by grey shading are regarded as "optimal" map-configurations.

Table 3.

Input A1	1: No Scaling	2: Range Scaling	3: Variance Scaling	4: Log_{10} Scaling
Input A2	5: No Scaling	6: Range Scaling	7: Variance Scaling	8: Log_{10} Scaling
Input A3	9: No Scaling	10: Range Scaling	11: Variance Scaling	12: Log_{10} Scaling
Input A4	13: No Scaling	14: Range Scaling	15: Variance Scaling	16: Log_{10} Scaling

Note:
1) **A1**: H_2O, N_2, O_2, CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2 (all gases including moisture).
2) **A2**: CO_2, CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2 (excluding atmospheric gases and moisture).
3) **A3**: H_2, CH_4, C_2H_6, C_2H_4, C_2H_2 (gases from oil degradation)
4) **A4**: CO, H_2, CH_4, C_2H_6, C_2H_4, C_2H_2 (combustible gases)

From Table 3, it can be seen that the "optimal" maps are mostly resulting from data set that has been pre-processed using the "range" method, which scales each component gas according to its maximum and minimum values. An example of the "optimal" maps, the map-configuration 14, is illustrated in Fig. 2.

Note that the U-matrix illustration (Map 1) essentially depicts the inherent "structure" of the DGA data [10]. The colour-depth of a 64-level grey scale as displayed in the U-matrix map is used to represent the Euclidean distance

between one neuron to the other. Therefore, the revealed data "structure" can be easily identified by means of black-border surrounding a lighter-coloured region. All corresponding data vectors located within a bordered region are in fact similar to one another in the vector space. On the other hand, the colour-depth depicted in the component planes (Map 2 to Map 7) of dissolved gases represents the magnitude; the deeper the colour, the larger is the magnitude of each dissolved gas component.

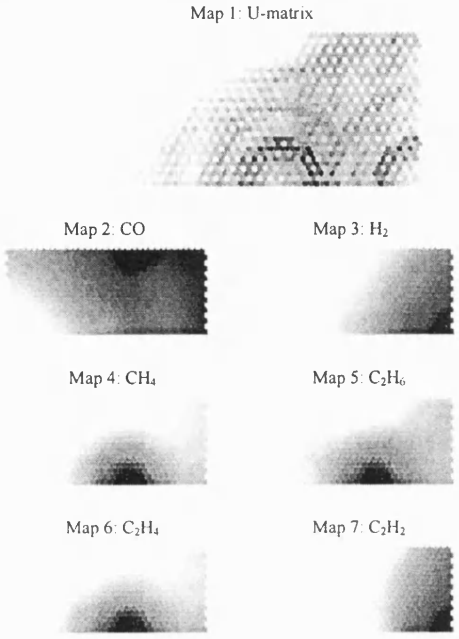


Fig 2. U-matrix and component planes for SOM configuration 14

As can be seen from Fig. 2, correlation is apparent for two categories of dissolved gases, viz. [H_2 C_2H_2] and [CH_4 C_2H_6 C_2H_4]. Owing to the observed relationship, two main groups of clusters are revealed in the U-matrix map. The bottom-middle clusters are formed by [CH_4 C_2H_6 C_2H_4] and the right-hand side clusters are formed by [H_2 C_2H_2]. These two groups of clusters also intercept with one another, forming a group of interception clusters between them.

C. Interpretation of Optimal SOM

The first step in interpreting the observed "structure" in the U-matrix map is to partition it into regions or clusters, as illustrated in Fig. 3. Detailed statistical analysis is then performed for each region, whereby the average concentration level (in part-per-million (PPM)) for each dissolved gas component is computed. The corresponding data subset for each region is obtained by performing BMU searches for weight-vectors in each region. The average concentration levels measured in PPM for dissolved gases are illustrated in Fig. 4.

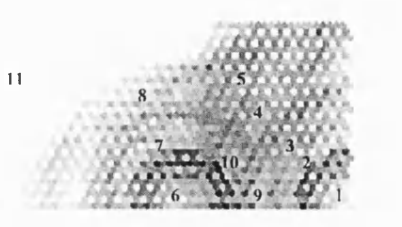


Fig. 3. Partition of hypothetical map into regions

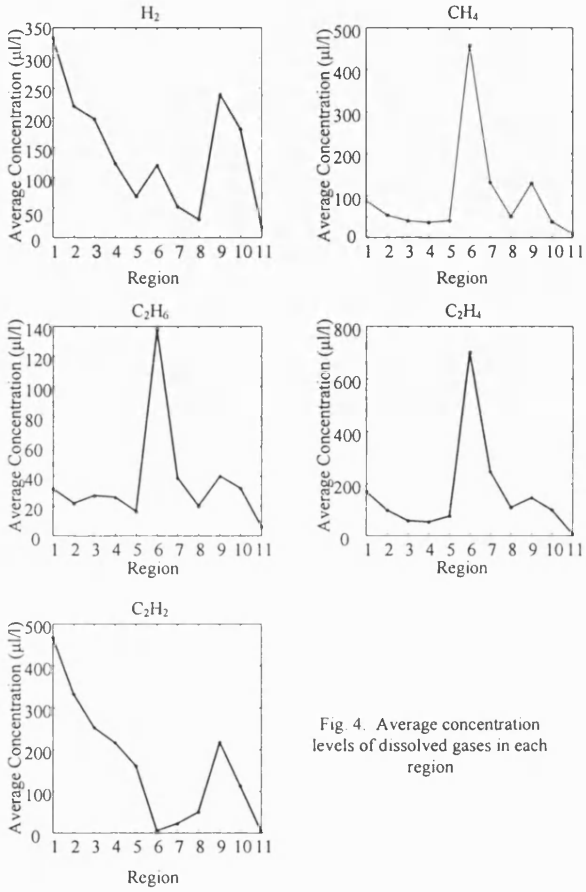


Fig. 4. Average concentration levels of dissolved gases in each region

Close correlation between the two categories of dissolved gases can also be observed from Fig. 4, as signified by the similarities in the distribution of average concentration levels. Both H_2 and specifically C_2H_2 are the dominant gases in regions 1 to 5, as observed from Fig. 4. Conversely, the CH_4, C_2H_6 and specifically C_2H_4, are found to be dominant in regions 6, 7 and 8.

As for the interception regions 9 and 10, although high levels of H_2 and C_2H_2 are found, reasonably high levels of CH_4, C_2H_6 and C_2H_4 are also observed. Interestingly, as compared with the [H_2 C_2H_2] dominant region, the average

levels of CH_4, C_2H_6 and C_2H_4 in the interception regions are found to be higher. Conversely, if compared with the [CH_4 C_2H_6 C_2H_4] dominant regions, the levels of H_2 and C_2H_2 are found to be higher in the interception regions.

As mentioned in the "Introduction", large amounts of H_2 and C_2H_2 are normally generated from electrical discharges (ED), and, high quantities of CH_4, C_2H_6 and C_2H_4 are normally liberated during the onset of a thermal fault (TF). In view of these observations, coupled with discoveries in detailed statistical analysis of each region, the "hypothetical" association of the type and severity of fault to each cluster can actually be established. Fig. 5 illustrates such an association. Note that the "ED_TF" indicates condition where there is simultaneous occurrence of ED and TF. The "dotted-arrows" indicate an increase in severity for a movement from outer to inner clusters.

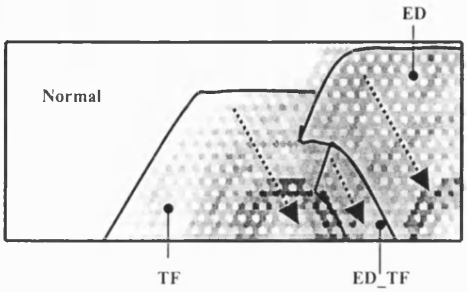


Fig. 5. Hypothetical assignment of fault regions

D. Comparison and Validation of Hypothesis

The hypothetical map as shown in Fig. 5 is compared with four previously mentioned DGA interpretation schemes. This is done so as to investigate the benefits of such a map over the conventional schemes and to ensure that such hypothesis is "valid" in the context of conventional DGA schemes. Comparison is performed by "overlapping" the interpretation and diagnoses as provided by these schemes onto the U-matrix map (see Map 1 in Fig. 2). The data set submitted for interpretation is actually weight-vectors of the optimal map configuration. Colour labelling is used to differentiate various identified faults or conditions. The interpreted maps are illustrated in Fig. 6.

As can be observed from Fig. 6, the interpretation and diagnoses as provided by DGA interpretation schemes are fairly consistent with the hypothetical map illustrated in Fig. 5. The location and the shape of Normal, TF and ED fault regions correspond quite closely with the hypothetical fault regions. The Dörnenburg's method [2] is, however, unable to interpret a large TF region. In addition, if guidelines provided by Dörnenburg are to be strictly followed, then no interpretation will result for a large ED region. On the other hand, although both Rogers Ratios [3] and Duval Triangle [4] are able to interpret correctly most

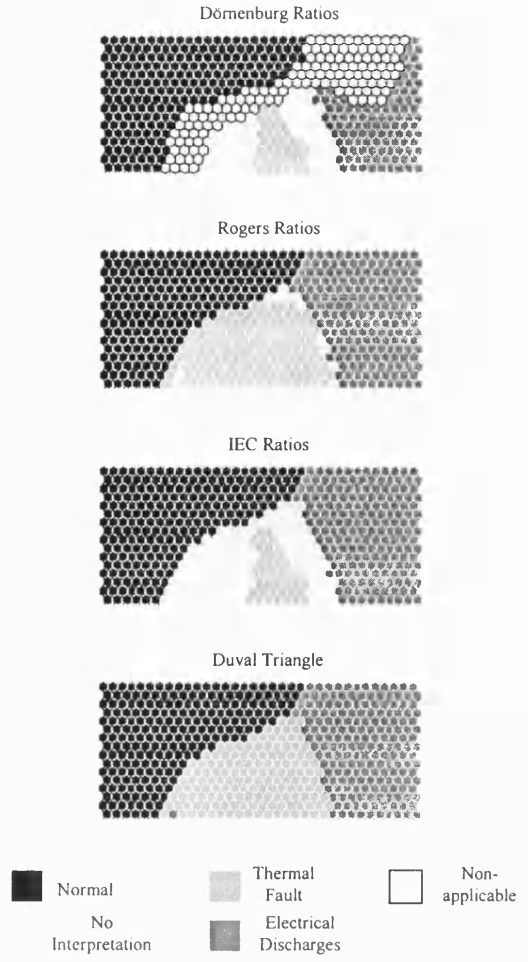


Fig. 6. Interpretation by DGA interpretation schemes

of the fault regions, they are unable to detect the simultaneous occurrence of faults, and thus are less effective compared to the hypothetical map illustrated in Fig. 5. Lastly, the IEC's method [1] is found to be less sensitive towards the less severe TF region.

In comparison, the hypothetical map, as illustrated in Fig.5, is more advantageous since the locations of fault regions are more realistically displayed and the simultaneous onset of faults can be detected. In addition, the severity of faults can be detected by observing the movement of DGA trajectories. If the DGA trajectory is moving from outer to inner clusters, then an increase in severity of a fault is suspected, and vice versa.

The hypothesis is validated by using real, observed fault cases. Fig. 7 illustrates some of these cases. As can be seen from Map 1 of Fig. 7, the DGA trajectory of that particular transformer does not venture into any of the fault region,

hence no fault is suspected. This diagnosis is in fact consistent with the observation by the DGA expert in the company, who attributes the changes in gas levels to the normal ageing of oil. In Map 2, it is clearly shown that the DGA trajectory has moved into the most severe TF region. This observation has been confirmed by DGA expert in the company, where a known thermal problem, peaking in year 1991 to 1992, actually happened in that transformer. Lastly, Map 3 illustrates a movement of DGA trajectory into the most severe ED region, as evident by the fact that arcing (or sparking) was known to have occurred at the clamping plates of that transformer.

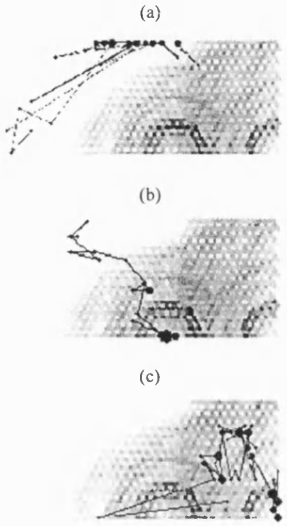


Fig. 7. Validation using real fault-cases

IV. CONCLUSIONS

It can be concluded that the proposed approach has many important advantages over traditional methods for analysis and interpretation of DGA data for the following reasons:

1) It does not depend on any actual fault cases and DGA schemes for its modelling, hence it is easy and cost-effective to implement.
2) It allows not only diagnosis of a fault based on visualisation of inherent data characteristics, but it also allows the CM of power transformers via the means of DGA trajectory plotting.
3) It offers a more consistent and convincing diagnosis as the revealed "structure" actually originates from within the real measured DGA records.

V. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions, in term of financial and physical resources, from the National Grid Company plc.

VI. REFERENCES

[1] IEC Standard Publication 599, "Interpretation of the analysis of gases in transformers and other oil-filled electrical equipment in service", International Electrotechnical Commission, 1999.
[2] E. Dörnenburg, W. Strittmatter, "Monitoring oil cooled transformers by gas analysis", Brown Boveri Review, vol. 61, no. 5, May 1974, pp. 238-247.
[3] R. R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis", 'IEEE Trans. Electrical Insulation, vol. E1-13, no. 5, Oct. 1978, pp. 349-354.
[4] M. Duval, "Fault gases formed in oil-filled breathing E.H.V. power transformers – the interpretation of gas analysis data", IEEE-PES Conference, paper C74-476-8.
[5] S. K. Bhattacharyya, R. E. Smith, T. A. Haskew, "A neural network based approach to transformer fault diagnosis using dissolved gas analysis data", Proceedings of the 1993 NAPS, Oct. 1993, pp. 125-129.
[6] Y. Zhang, X. Ding, Y. Liu, P. J. Griffin, "An artificial neural network based approach to transformer fault diagnosis", IEEE Trans. Power Delivery, vol. 11, no. 4, Oct. 1996, pp. 1836-1841.
[7] C. E. Lin, J. M. Ling, C. L. Huang, "An expert system for transformer fault diagnosis using dissolved gas analysis", IEEE Trans. Power Delivery, vol. 8, no. 1, Jan. 1993, pp. 231-238.
[8] K. Tomsovic, M. Tapper , T. Ingvarsson, "Fuzzy information approach to intergrating different transformer diagnosis methods", IEEE Trans. Power Delivery, vol. 8, no. 3, Jul. 1993, pp. 1638-1645.
[9] T. Kohonen, Self-Organising Maps, 2nd edition, Springer-Verlag, 1997.
[10] A. Ultsch, H. P. Simon, "Kohonen's self-organising feature maps for exploratory data analysis", International Neural Network Conference, 1990, pp. 305-308.

VII. BIOGRAPHIES

K. F. Thang is a PhD student in the Power and Energy System Group at the University of Bath, England. His research interests are the AI applications to power systems and data-mining for condition monitoring of power transformers. He is an associate member of the IEE

R. K. Aggarwal received the B. Eng and PhD degrees from the University of Liverpool, England, in 1970 and 1973, respectively. He joined the Power and Energy System Group at the University of Bath, England, where he is now a Professor. His main research interests are power system modelling and application of digital-technology and AI to protection and control. He has published over 200 technical papers and co-authored three text books. Prof. Aggarwal is a senior member of the IEEE and fellow of the IEE.

A. J. McGrail is a transformer and condition monitoring specialist with the National Grid Company plc. in the UK. His work involves both research and practical engineering, using both traditional analysis and AI/ANN tools to extract as much information as possible from available data. He has a degree in Physics, a Masters and PhD in Electrical Engineering and an MBA. He is a member of the IEEE.

D. G. Esp is a Research Engineer with the National Grid Company plc. in the UK as well as the pre-privatisation Central Electricity Generating Board (CEGB) since 1980. His interests have included high-integrity computing, AI applications for power systems, data mining and web technology. In recent years he has applied such technologies to areas such as alarm handling, demand forcasting and condition monitoring, in particular as regards the interpretation of high voltage transformer DGA. He is a member of the IEE.

Data Mining Approach for Analysis of Power Transformer Dissolved Gas Records Using the Self-Organising Map

K. F. Thang, *Student Member, IEEE*, R. K. Aggarwal, *Senior Member, IEEE*,
A. J. McGrail, *Member, IEEE*, D. G. Esp

Abstract--**Incipient faults in power transformers can degrade the oil and cellulose insulation, leading to the formation of dissolved gases. Although interpretation schemes for dissolved gases are readily available, which relate dissolved gas records to condition of power transformers, it is pointed out in this paper that these methods contain some inherent weaknesses. In view of that, this paper introduces a novel approach for analysis of dissolved gas records, which can lead to a more convincing interpretation and accurate diagnosis. The proposed approach, which is based on data mining methodology and the application of self-organising map, has been validated using real fault-cases and thereby is proven to be capable of addressing the inherent weaknesses of the conventional interpretation methods.**

Index Terms--**Data mining, dissolved gases, incipient faults, power transformers, self-organising map.**

I. Introduction

INCIPIENT faults in power transformers, either electrical or thermal in nature, can degrade the oil and cellulose insulation, leading to the formation of gases that dissolve in oil. Such faults can be detected and monitored using dissolved gas analysis (DGA).

DGA requires the identification and quantification of seven major dissolved gases, as stated in the IEC 60599 Standard [1]. These gases are carbon dioxide (CO_2), carbon monoxide (CO), hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4) and acetylene (C_2H_2). Composition of dissolved gases is closely linked to the type and severity of a fault [2]. For example, in the case of a fault with low-level energy such as partial discharges (PD), the main gas produced is H_2. In other types of fault, degradation of oil is mainly caused by heat. Therefore, oil degradation at normal operating temperature produces mainly H_2 and CH_4. A higher fault temperature, resulting from hot-spots or overheating, produces mainly CH_4. In addition, C_2H_6 and C_2H_4 will be produced in increasing quantity if the fault temperature continues to rise. In the case of much higher temperature resulting from disruptive fault such as electrical discharges (ED), the production of C_2H_2 becomes significant. Furthermore, if cellulose materials such as paper and pressboard etc., are involved in the immediate vicinity of faults, both CO_2 and CO will also be generated.

Existing DGA interpretation methods relate the composition of dissolved gases to health and condition of power transformers. Although widely adopted by power utilities, these schemes are identified to be less than perfect. Even though attempts have been made to utilise AI techniques to enhance the capability of interpretation, as discussed in this paper, these approaches still contain some intrinsic limitations.

Comparative studies of conventional and AI approaches for DGA interpretation are presented in the first part of the paper. The second part of the paper concentrates on introducing the novel approach that is based on data mining (DM) methodology and the use of self-organising map (SOM) for analysis and interpretation of DGA data.

II. Comparative Studies of Conventional and AI Approaches for DGA Interpretation

Comparative studies of conventional and AI approaches for DGA interpretation have been conducted. The studies focus on evaluating the differences and similarities among various approaches and are presented as follow:

A. Conventional Approaches

Several renowned DGA interpretation methods are, for example, Dörnenburg Ratios [3], Rogers Ratios [2], Duval's Triangle [4] and the IEC Ratios [1]. These methods have been implemented, either in modified or improvised format, by various power utilities. The implementation of these methods requires the computation of several key gas-ratios, as shown in Table I. Fault diagnosis is accomplished by associating the values of these ratios with a set of pre-defined conditions of power transformers.

Generally, detection of two distinctive types of fault is possible, i.e. electrical fault (EF) and thermal fault (TF). EF can be further divided into partial discharges (PD) and electrical discharges (ED). Before subjecting DGA records to interpretation, a decision has to be made on whether fault

K. F. Thang is with the Department of Electronic and Electrical Engineering, University of Bath, Bath, BA2 4AY, UK (e-mail: eepkft@bath.ac.uk).

R. K. Aggarwal is with the Department of Electronic and Electrical Engineering, University of Bath, Bath, BA2 4AY, UK (e-mail: eesrka@bath.ac.uk).

A. J. McGrail is with the National Grid Company plc, Leatherhead, KT22 7ST, UK (e-mail: Tony.McGrail@uk.ngrid.com).

D. G. Esp is with the National Grid Company plc, Wokingham, RG31 5BN, UK (e-mail: David.Esp@uk.ngrid.com).

diagnosis is necessary based on a comparison of dissolved gas concentrations with a set of "benchmark" values, which are also known as the typical value of gas concentrations. If all measured dissolved gas concentrations are below the typical values, then the corresponding transformer can be regarded as operating normally. If a large historical DGA database is available, the "benchmark" values should be calculated from it, based on the 90% or 95% limits [1].

TABLE I
KEY GAS-RATIOS OF CONVENTIONAL DGA INTERPRETATION SCHEMES

Interpretation Scheme	Key Gas-Ratios
Dörnenburg Ratios	$\frac{CH_4}{H_2}$; $\frac{C_2H_2}{C_2H_4}$; $\frac{C_2H_6}{C_2H_2}$; $\frac{C_2H_2}{CH_4}$
Rogers Ratios	$\frac{CH_4}{H_2}$; $\frac{C_2H_6}{CH_4}$; $\frac{C_2H_4}{C_2H_6}$; $\frac{C_2H_2}{C_2H_4}$
Duval's Triangle *	$\%C_2H_2$; $\%C_2H_4$; $\%CH_4$
IEC Ratios	$\frac{C_2H_2}{C_2H_4}$; $\frac{CH_4}{H_2}$; $\frac{C_2H_4}{C_2H_6}$

Note: For Duval's Triangle, %[Gas] is calculated by using [Gas] as numerator and the sum of three gases as denominator; percentage is then computed on the ratio.

Although well received by power utilities, the foregoing interpretation methods are known to have some weaknesses:

1) Key gas-ratios and their limits are mainly determined based on judgement and experience of experts; there is no systematic attempt on high-level analysis of real DGA data.
2) Interpretation is sometimes impossible due to the inability to cover all possible combinations of ratio values, whereby the diagnosis has to be dependent on experts.
3) Inconsistency and ambiguity in implementation due to different interpretations by various schemes.
4) Inability to detect with high confidence faults that occur simultaneously within the transformer.

B. AI Approaches

Attempts have been made to utilise AI techniques to tackle the weaknesses of conventional interpretation methods. These AI approaches can be classified into two categories.

The first category of AI approaches relies on the use of supervised artificial neural networks (ANNs) in an attempt to establish the intrinsic relationship between the actual DGA records and corresponding condition of transformers, as reported in [5]-[7]. The condition of transformers can be obtained either through inspection on suspected faulty transformers [5]-[6] or the use of conventional interpretation schemes [7]. A set of input-target output samples is required for the training of supervised ANN. The inputs are, for example, dissolved gases [5]-[6] or the values of the key gas-ratios [7]; the target-outputs are the corresponding condition of transformers.

The second category of AI approaches integrates both fuzzy logic (FL) and expert system (ES), as reported in [8]-[9]. This approach allows the incorporation of expert experience into its knowledge-base, whereas the core is mainly built on several conventional interpretation schemes. The incorporation of FL is beneficial since the uncertainty on normality thresholds and ratio values can be effectively dealt-with.

While improvement over conventional methods have been reported, drawbacks of these AI approaches are identified as follows:

1) Acquisition of DGA records with observed faults in corresponding transformers is difficult since it is costly and impractical to disconnect and dismantle a particular transformer for the purpose of confirming a suspected fault. Therefore, supervised ANNs would not be able to generalise due to a lack of "good" fault-cases for the training.
2) The dependency of fuzzy expert system (FES) on conventional interpretation schemes for its modelling has led to the inheritance of some of the inherent weaknesses of these schemes.

III. THE DATA MINING APPROACH

Data mining (DM) refers to a process whereby various techniques, either statistical or AI-based, are employed to unearth or "mine" valuable information from the data. It is known that DGA is always performed on power transformers as part of the condition monitoring (CM) procedures. Eventually, an enormous amount of DGA records are gathered and stored in the database. These records actually contain valuable information regarding the series of "events" or "incidence" (e.g. incipient faults) that have taken place within the transformers. If the hidden information can be extracted, comprehended and visualised, our understanding on the condition and health of transformers can be improved. Thus, this paper explores the potential of the DM approach for analysis and interpretation of DGA data.

A. Introduction to Self-Organising Map: A DM Technique

The self-organising map (SOM) is constructed by neurons located on a regular two-dimensional grid, as illustrated in Fig. 1 [10]. The configuration of SOM is such that it defines a mapping from the input space $\Re^n$, where n is the dimension of input space onto a two-dimensional array of neurons. Every neuron i is associated with an n-dimensional weight vector $m_i = [\mu_{i1}, \mu_{i2}, \mu_{i3}, \ldots, \mu_{in}]^T$. An input vector $x = [\xi_1, \xi_2, \xi_3, \ldots, \xi_n]^T$ is connected to all neurons in parallel via weight values μ_{ij}, where $j = 1, \ldots, n$, which are different for every neuron.

Once the weight vectors m_i have been initialised, the training of SOM commences by first choosing an input vector x, at time-step t, randomly from the input data. The Euclidean distances between x and all weight vectors m_i are calculated. The weight vector that is closest to x is known as the best-matching unit (BMU), which is denoted as c. Equation (1) illustrates the above stated relation. All weight vectors are

then updated according to (2). Note that $h_{ci}(t)$ is the neighbourhood kernel around the BMU c, at time t. It can be defined as the function of the learning rate, $\alpha(t)$, and the neighbourhood function, $h(\|r_c - r_i\|, t)$, as illustrated in (3). The rectangular function, which is illustrated in Fig. 2, is considered to be a suitable neighbourhood function for the analysis of DGA data.

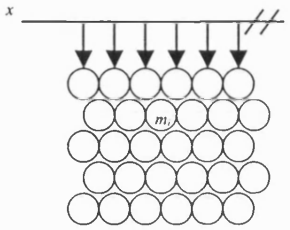


Fig. 1. Neurons arranged in hexagonal lattice in a two-dimensional SOM.

$$c = \arg\min_i \{\|x - m_i\|\} \tag{1}$$

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)] \tag{2}$$

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c - r_i\|, t) \tag{3}$$

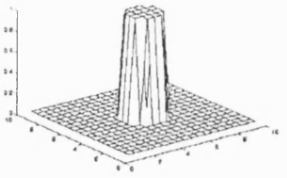


Fig. 2. The rectangular function.

If the SOM is sufficiently trained in accordance with some preset criterion, effective mapping from the input space to the two-dimensional SOM can thus be achieved. The hidden information of input data can then be visualised based on the trained SOM. There are two criteria in which a SOM mapping can be considered as effective:

1) The weight vectors of SOM must be similar to that of input vectors for effective approximation, i.e. short Euclidean distances between input vectors and their BMUs.
2) Optimal configuration of neurons on the SOM for effective visualisation of intrinsic data characteristics.

The optimal SOM is selected from many trials based on these two aspects. In fact, the selection process is founded on the calculation of quantisation error (QE), which measures the first criteria, and topographic error (TE), which measures the second criteria as described above.

B. Analysis of DGA Data by Using SOM

A set of DGA data was compiled for the feasibility study of the proposed DM approach. This data set, which comprises 755 actual DGA records, corresponds to 240MVA transmission transformers from three different manufacturers and of two distinct voltage ratios. The subsets of this DGA data are shown in Table II. Note that in all instances, the corresponding oil was sampled from the bottom of the transformer tanks.

TABLE II
SUBSETS OF DGA DATA

Manufacturer *	Voltage Ratio	Power Rating (MVA)	Number of Records
A	400/132	240	72
A	275/132	240	91
B	400/132	240	94
B	275/132	240	187
C	400/132	240	108
C	275/132	240	203

Note: Characters are used to represent manufacturers

The DGA data was pre-processed and arranged into four sets of inputs with different logical combination of the ten recorded dissolved gases, as shown in Table III. This is done so as to determine the key or important dissolved gases by visualising the optimal SOMs that have been trained based on these input sets.

TABLE III
FOUR DGA INPUT SETS

Data	Dissolved Gases	Description
Set A	H_2O; N_2; O_2; CO_2; CO; H_2; CH_4; C_2H_6; C_2H_4; C_2H_2	All recorded dissolved gases including moisture
Set B	CO_2; CO; H_2; CH_4; C_2H_6; C_2H_4; C_2H_2	Dissolved gases from cellulose and oil degradation
Set C	CO; H_2; CH_4; C_2H_6; C_2H_4; C_2H_2	Combustible dissolved gases
Set D	H_2; CH_4; C_2H_6; C_2H_4; C_2H_2	Dissolved gases from oil degradation

Following the completion of training, an optimum SOM is selected for each set of inputs. It is found that regardless of the number of dissolved gases included in each set of inputs, two obvious characteristics that are contributed by H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 are always observed. Examples of the optimal SOM for Sets B and C are illustrated in Figs. 3 and 4 respectively. Note that the optimal SOM is visualised based on the view of component-planes. A 64-level grey-scale is

used to depict the concentration (measured in micro-litre per litre, $\mu l/l$) of each dissolved gas; the lighter the colour, the higher is the concentration of a particular dissolved gas.

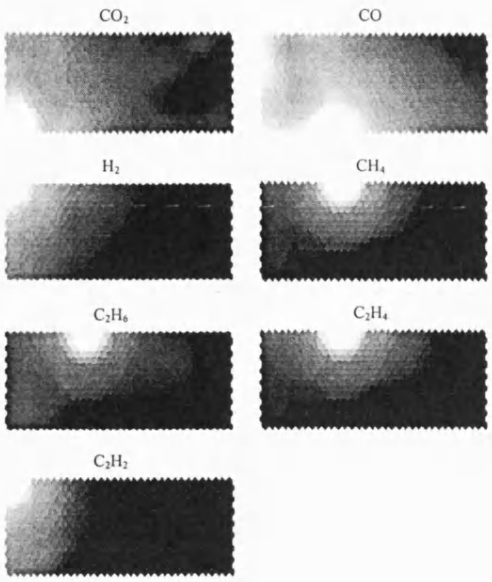


Fig. 3. Component-planes of optimal SOM for Set B.

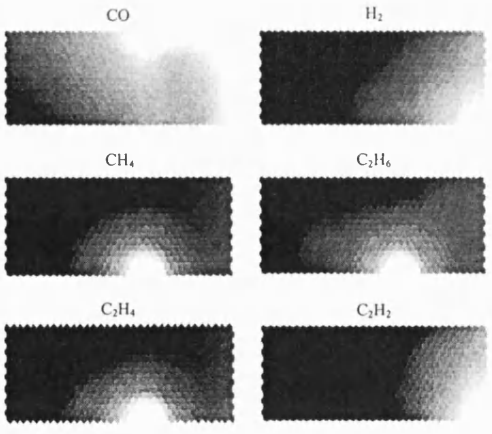


Fig. 4. Component-planes of optimal SOM for Set C.

As illustrated in Figs. 3 and 4, similarity in inherent characteristic is clearly observed between H_2 and C_2H_2, and among CH_4, C_2H_6 and C_2H_4 as well. In fact, the observed correlation in these two groups of dissolved gases is comprehensible.

For example, a large quantity of C_2H_2 will be formed during the onset and evolution of electrical discharges (ED). Moreover, H_2 will also be generated in large amounts due to extremely high temperatures associated with the fault. Thus, the correlation between H_2 and C_2H_2 is logical and is successfully unearthed by SOM. On the other hand, in case of thermal fault (TF), large amounts of CH_4, C_2H_6 and C_2H_4 will be produced. However, the concentration of C_2H_2 is not significant due to the lower fault temperature. Therefore, the SOM has also succeeded in learning and displaying this fact. In addition, if the afore-mentioned faults also involve cellulose insulation, the generation of CO_2 and CO becomes significant. That is why areas of high concentration of H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2 also correspond to quite high concentration of CO_2 and CO.

C. Interpretation of Optimal SOM

A more effective way of visualising the overall inherent characteristics of DGA data is the view of "u-matrix" map; it is simply a matrix of Euclidean distances among all neurons of SOM. It is used to gain a global view of clusters resulting from the correlation among dissolved gases. As an example, the "u-matrix" of optimum SOM for Set C is illustrated in Fig. 5.

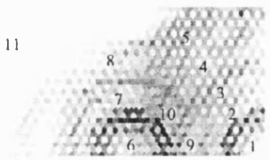


Fig. 5. The "u-matrix" of optimal SOM for Set C.

The first step in interpreting the observed inherent characteristics in the "u-matrix" map is to partition it into 11 regions, as illustrated in Fig. 5. Detailed statistical analysis is then performed on each region, whereby the average concentration of dissolved gases is computed. The corresponding input vectors for each region are gathered by performing BMU searches for neurons. The average concentration for every key dissolved gas across various regions is as illustrated in Fig. 6.

Close correlation among two groups of dissolved gases is revealed in Fig. 6, as signified by the resemblance of the concentration curves. Moreover, it is proven that both H_2 and C_2H_2 are the dominant gases in regions 1 to 5. In contrast, CH_4, C_2H_6 and C_2H_4 are found to be dominant in regions 6 to 8. Moreover, considerable concentration of all key dissolved gases is observed in regions 9 and 10, since they form the interception regions between the two main clusters.

It is thus obvious from the foregoing that the inherent characteristics of DGA data as illustrated in the "u-matrix" map can be associated with the condition and health of power

transformers, as sufficiently proven by the common sense knowledge in dissolved gas formation in relation to faults and consistent revelations in statistical analysis. The interpretation of the "u-matrix" map in relation to faults is as shown in Fig. 7, which illustrates that the diagnosis and monitoring of four conditions is possible, i.e. ED, TF, simultaneous occurrence of ED and TF, and the normal operating condition where no incipient fault is suspected.

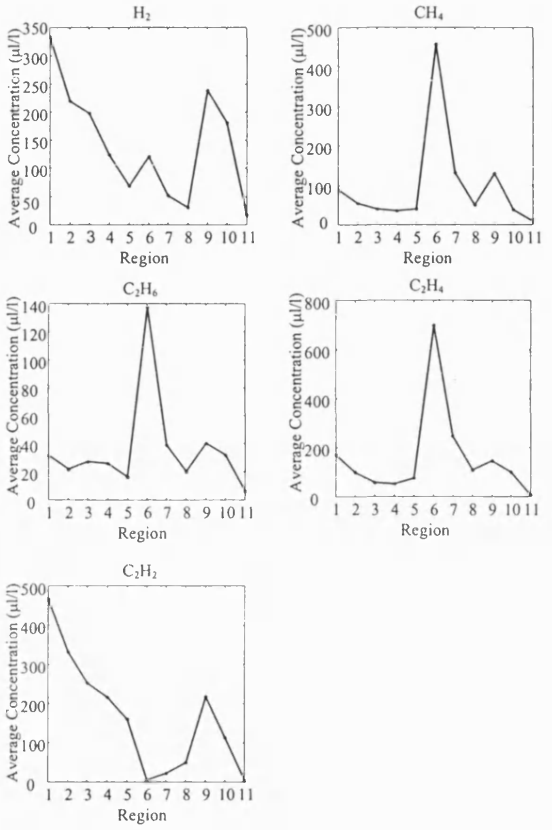


Fig. 6. Average concentration of dissolved gases across various regions.

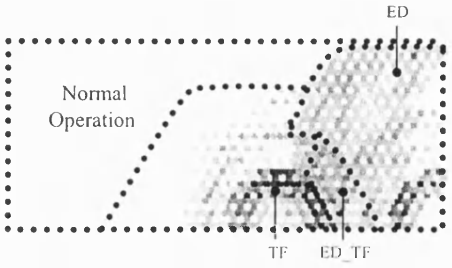


Fig. 7. Interpretation of "u-matrix" map in relation to faults.

D. Validation of Extracted Knowledge

The "extracted" knowledge with regard to the association of inherent DGA characteristics with the health and condition of power transformers is further validated using real, confirmed fault cases. Fig. 8 depicts some examples of these cases.

As can be observed from Fig. 8(a), the DGA trajectory of that particular transformer does not venture into any of the identified fault regions, hence no incipient fault is suspected. This diagnosis is in fact consistent with the observation of an expert in NGC, who attributes the changes in gas levels to the normal ageing of oil. In Fig. 8(b), it is clearly depicted that the DGA trajectory has slowly moved into the identified TF region. This observation has also been confirmed by the NGC's expert, and is directly attributed to a known transformer thermal problem, peaking in year 1991 to 1992. Lastly, Fig. 8(c) illustrates the movement of DGA trajectory into the identified ED region, as evident by the fact that arcing (or sparking) was known to have occurred at the clamping plate of that transformer.

Hence, the interpretation on transformer condition as given by the proposed approach matches the actual condition and fault incidence as observed by a transformer expert.

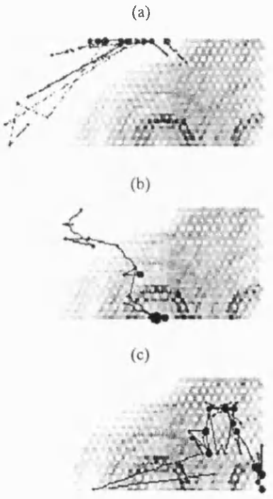


Fig. 8. DGA trajectories of power transformers.

IV. Conclusions

In conclusion, the proposed DM approach for analysis and interpretation of DGA data using the SOM has successfully resolved several inherent weaknesses of the conventional DGA interpretation schemes, as listed below:

1) Determination of gas-ratios and their limits is not necessary; the modelling process is based entirely on recorded dissolved gas records.

2) Diagnosis and interpretation of the health and condition of power transformers are always guaranteed; only actual dissolved gas records are required for the interpretation.
3) Uncertainty and ambiguity in interpretation is avoided owing to the fact that the interpretation is founded on the inherent characteristics of DGA data.
4) Visualisation on transformer condition is now possible via the means of DGA trajectory; onset of an incipient fault can therefore be effectively detected and monitored "visually".
5) It is relatively cost-effective to implement the proposed approach since only historical DGA database is required for its modelling.

Finally, the proposed approach can be used as a decision support tool by transformer engineers in addition to conventional approaches for the monitoring of health and incipient faults of power transformers.

V. Acknowledgment

The authors gratefully acknowledge the contributions, in terms of financial support and data resources, of the National Grid Company plc., UK.

VI. References

[1] *Interpretation of the Analysis of Gases in Transformers and Other Oil-Filled Electrical Equipment in Service*, IEC Standard Publication 599, 1999.
[2] R. R. Rogers, "IEEE and IEC codes to interpret incipient faults in transformers, using gas in oil analysis", *IEEE Trans. Electrical Insulation*, vol. EI-13, no. 5, pp. 349-354, Oct. 1978.
[3] E. Dörnenburg and W. Strittmatter, "Monitoring oil-cooled transformers by gas analysis", *Brown Boveri Review*, vol. 61, no. 5, pp. 238-247, May 1974.
[4] M. Duval, "Fault gases formed in oil-filled breathing E.H.V. power transformers – the interpretation of gas analysis data," in *Proc. IEEE Power Engineering Society Conf.*, paper C74-476-8.
[5] S. K. Bhattacharyya, R. E. Smith and T. A. Haskew, "A neural network approach to transformer fault diagnosis using dissolved gas analysis data," in *Proc. North American Power Symposium (NAPS)*, pp. 125-129, Oct. 1993.
[6] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An artificial neural network based approach to transformer fault diagnosis", *IEEE Trans. Power Delivery*, vol. 11, no. 4, pp. 1836-1841, Oct. 1996.
[7] J.L . Guardado, J.L . Naredo, P. Moreno and C. R. Fuerte, "A comparative study of neural network efficiency in power transformers diagnosis using dissolved gas analysis", *IEEE Trans. Power Delivery*, vol. 16, no. 4, pp. 643-647, Oct 2001.
[8] C. E. Lin, J. M. Ling and C. L. Huang, "An expert system for transformer fault diagnosis using dissolved gas analysis", *IEEE Trans. Power Delivery*, vol. 8, no. 1, pp. 231-238, Jan 1993.
[9] K. Tomsovic, M. Tapper and T. Ingvarsson, "Fuzzy information approach to integrating different transformer diagnosis methods", *IEEE Trans. Power Delivery*, vol. 8, no. 3, pp. 1638-1654, July 1993.
[10] T. Kohonen, *Self-Organising Maps*, 2nd edition, Springer-Verlag, 1997.

VII. Biographies

K. F. Thang (Student Member, IEEE) received the B.Eng. (Hons.) from the University of Bath in 1998. He is a now a Ph.D. student of the Power and Energy System Group at the Department of Electronic & Electrical Engineering at the same university. His research interests are the AI application to various aspects of power systems and data mining approaches for condition monitoring of power transformers.

R. K. Aggarwal (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees from the University of Liverpool, England, in 1970 and 1973, respectively. He joined the Electrical Power and Energy Systems Group at the University of Bath, England, where he is now a Professor and Head of the Electrical Power and Energy Systems Group. His main research interests are power system modelling and application of digital-technology and AI to protection and control. He has published over 300 technical papers and co-authored four textbooks.

A. J. McGrail (Member, IEEE) is a transformer and condition-monitoring specialist with the National Grid Company plc. in the UK. His work involves both research and practical engineering, using both traditional analysis and AI/ANN tools to extract as much information as possible from available data. He has a degree in Physics, a Masters and Ph.D. in Electrical Engineering and an MBA.

D. G. Esp is a Research Engineer with the National Grid Company plc. in the UK as well as the pre-privatisation Central Electricity Generating Board (CEGB) since 1980. His interests have included high-integrity computing, AI applications for power systems, data mining and web technology. In recent years he has applied such technologies to areas such as alarm handling, demand forecasting and condition monitoring, in particular as regards the interpretation of high voltage transformer DGA.